

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 30, 2004, 06:15:28 ; Search time 5051 Seconds
(without alignments)
3192.593 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796
Sequence: 1 MPRPFHKLIFSSTIOEKRL.....LLRTRDFVLKVTAPRVNEYV 341

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh
-O/cgn2_1/USPTO spool/US10088187/runat_29122004_115842_119/app.query.fasta_1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10088187 @CNC 1.1 3731 @runat_29122004_115842_119 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1796	100.0	1026	CQ805406 Sequence
2	1796	100.0	1495	AX101036 Sequence
3	1796	100.0	1538	AF289052 Arabidops
4	1781	99.2	1495	AX101038 Sequence

	5	1768	98.4	1494	6	AX101039	Sequence
	6	1614.5	89.9	5000	8	AF289051	Arabidops
C	7	1614.5	89.9	79186	8	AP000735	Arabidops
	8	1601.5	89.2	5000	6	AX101027	Sequence
	9	1513	84.2	990	8	AY517929	Brassica
	10	1509	84.0	990	8	AY356368	Brassica
	11	1113.5	62.0	1146	8	AY042868	Arabidops
	12	1098	61.1	125021	8	AC007504	Arabidops
	13	946	52.7	681	6	AX505578	Sequence
	14	850	47.3	78589	8	AP006145	Lotus cor
C	15	808	45.0	243777	2	AC137825	Medicago
	16	359	20.0	122693	8	AC145449	Medicago
	17	357	19.9	114098	8	AC144459	Medicago
	18	333	18.5	48368	2	AC140105	Medicago
	19	332.5	18.5	92281	8	AP004896	Lotus cor
C	20	301.5	16.8	48368	2	AC140105	Medicago
	21	298.5	16.6	1361	6	AX558824	Sequence
	22	293.5	16.3	1041	6	AX558782	Sequence
	23	281	15.6	1762	8	AK101941	Oryza sat
	24	281	15.6	1884	8	AK067571	Oryza sat
	25	270	15.0	1469	8	AK108660	Oryza sat
	26	249.5	13.9	129647	8	AC097280	Oryza sat
C	27	249	13.9	108028	8	AC135233	Medicago
	28	239	13.3	1014	8	ATH441117	Arabidops
C	29	237.5	13.2	129204	8	AC146747	Medicago
	30	236.5	13.2	1409	8	AK063506	Oryza sat
	31	236	13.1	132741	8	CNS08CAA	Arabidops
	32	229.5	12.8	2000	6	AX508278	Sequence
	33	229	12.8	573	8	BT010604	Arabidops
	34	220	12.2	204120	8	AY661659	Sorghum b
	35	217	12.1	661	8	BT008541	Arabidops
	36	217	12.1	809	8	BT006139	Arabidops
C	37	202	11.2	1495	8	AK109729	Oryza sat
	38	195.5	10.9	937	8	BT005553	Arabidops
	39	195.5	10.9	1113	8	BT004099	Arabidops
	40	195.5	10.9	1221	8	AY087853	Arabidops
C	41	192.5	10.7	90596	8	TI5B16	Arabidops
	42	192.5	10.7	198697	8	ATCHRIV4	Arabidops
	43	191.5	10.7	140383	8	CNS08CD5	Arabidops
C	44	188.5	10.5	125021	8	AC007504	Arabidops
C	45	185.5	10.3	36495	8	AB015473	Arabidops

ALIGNMENTS

RESULT 1	CQ805406	Sequence 1817 from Patent WO2004035798.	1026 bp	DNA	linear	PAT 10-MAY-2004
LOCUS	CQ805406	Sequence 1817 from Patent WO2004035798.				
DEFINITION	CQ805406					
ACCESSION	CQ805406					
VERSION	CQ805406.1	GI:47111337				
KEYWORDS						
SOURCE	Arabidopsis thaliana (thale cress)					
ORGANISM	Arabidopsis thaliana					
REFERENCE	1	Inze, D., de Veylder, L. and Vlieghe, K.				
AUTHORS		Identification of novel e2f target genes and use thereof				
TITLE		Patent: WO 2004035798-A 1817 29-APR-2004;				
JOURNAL		CropDesign N.V. (BE)				
FEATURES		Location/Qualifiers				
source		1..1026				
		/organism="Arabidopsis thaliana"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:3702"				

Alignment Scores:	1.71e-139	Length:	1026
Pred. No.:	1796.00	Matches:	341
Score:	100.00%	Conservative:	0

Best Local Similarity: 100.00%		Mismatches: 0	
Query Match: 100.00%		Indels: 0	
DB: 6		Gaps: 0	
US-10-088-187A-11 (1-341) x CQ805406 (1-1026)			
QY	1	MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu	20
DB	1	ATGCCACGCCCTTCTTCATAGTTGATTTCTCATCCACTATCCAGAAAAACGCTG	60
QY	21	ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu	40
DB	61	AGGGTCCAGATAAGTTTGTGATTAATTCAGGATGAGCTTCGGTTCGTGTGCAC	120
QY	41	ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp	60
DB	121	ACAGTACCTGATGGTCATGTTTGGCGGTGAGGACTAAGGAAAGCTGACAA	180
QY	61	PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu	80
DB	181	TTTCAAGATGGTGGCAGAGTTTGTGACCGTTACTCCTTCGATTTGTTATCTTTG	240
QY	81	IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer	100
DB	241	ATTTTAGATATGAAGAAACTCTGCTTCAGCGTCTACATTTTCAATTTATCCACTCT	300
QY	101	GlulLeuAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla	120
DB	301	GAGATCAATTACATTCACCGGTCTCATGGATTCGCTCAACACCACTTCAAAACGCGCC	360
QY	121	ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr	140
DB	361	CGTTTGTGTAAGACCTTGAAGATGAAGATGCGGAGGTCACTTCTTCTCTGTGTAC	420
QY	141	ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle	160
DB	421	CCATCACCACTTCTCAGTCTACAGTACCAGCCCAACAAAGGATATGCTAGTTACGCCATC	480
QY	161	GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro	180
DB	481	CAAACTTGTCTACTGGACCACTTAAAGCTGAAGAGCCAAACCCCAAAAAATACCT	540
QY	181	LysLysArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro	200
DB	541	AAAAAGAGAGGAGGAGGAAGAAAAATGCTGATCTCTGAGGAAATAAATCACTCAGCTCCG	600
QY	201	ArgAspAspProGluLeuAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg	220
DB	601	CGAGATGATGATCCAGAAACCGTTCAAAGTTCTACGAGTGTCTTTCGGAGAAAGAGA	660
QY	221	ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaLysThrPheGluProThr	240
DB	661	ACCGTGACTCAGAGAAGAGAGAGAGAGCCATCAATGCAGCCAAACGTTCCGAACCAACA	720
QY	241	AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr	260
DB	721	AACCTTTCTTCAGAGTGTCTTCGCAACCATCTATCTATACAGAGTTGCATCATGAT	780
QY	261	LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln	280
DB	781	CTTCTTCTGGTGTGCTGAGAGTACCTAAGTGGGATCTCCGGGTTCATCAAGTCCAG	840
QY	281	LeuAlaGluLysGlnTrpProValArgCysLeuTyrIleAlaGlyArgAlaLysPheSer	300
DB	841	CTTCGGAGAAACAATGGCTGTTCGATGTCTCTACAAAGCCGGGAGAGCCAAATTCAGT	900
QY	301	GlnGlyTrpTyrGluPheThrLeuGluAsnAsnLeuGlyGlyAspValCysValPhe	320
DB	901	CAAGATGGTACGAATTCATCTAGAGAACCACTTAGGAGAGAGACGCTCTGTGTGTT	960
QY	321	GlulLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr	340
DB	961	GAGCTGCTCAGAACAGAGATTTGTTTTGAAAGTGACAGCCTTTCAGTCAACAGGTAC	1020
QY	341	Val 341	
DB	1021	GTC 1023	
RESULT 2			
LOCUS	AX101036	1495 bp	DNA linear PAT 10-APR-2001
DEFINITION	Sequence 10 from Patent WO0121822.		
ACCESSION	AX101036		
VERSION	AX101036.1	GI:13619892	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE	Dean, C. and Levy, Y. Y.		
AUTHORS	Methods and means for modification of plant flowering		
TITLE	characteristics		
JOURNAL	Patent: WO 0121822-A 10 29-MAR-2001;		
FEATURES	Plant Bioscience Limited (GB)		
source	Location/Qualifiers		
	1. .1495		
	/organism="Arabidopsis thaliana"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:3702"		
	269. .1294		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC36501.1"		
	/db_xref="GI:13619893"		
	/translation="MPRPFHKLIFSSTIQEKRLRVPDKFVKFDELSVAVALTVDP		
	GHVWVGLKADNKIWFQDQWQFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEI		
	NYHSTGLMDSAHNHFKARLFDELEDAEVIIPSSVYPSPLPSTVPANKYASSAI		
	QTLFTGPVKAEEPTPKIPKGRKKONADPEINSSAPDDDDPENRKFYFESASAR		
	KRTVTAERERAINAAKTFTPTNPFVRVLRPSLYRGCMILYPSGFAEYLSGISGF		
	IKVQLAEKQWPVRCLYKAGRAKFSQGWYEFTLENNLGEQDVCFELLRDFVLKVTA		
	FRVNEYV"		
ORIGIN			
Alignment Scores:	2.73e-139	Length:	1495
Pred. No.:	1796.00	Matches:	341
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	6	Gaps:	0
DB:			
US-10-088-187A-11 (1-341) x AX101036 (1-1495)			
QY	1	MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu	20
DB	269	ATGCCACGCCCTTCTTCATAGTTGATTTTCTCATCCACTATCCAGAAAAACGCTG	328
QY	21	ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu	40
DB	329	AGGGTCCAGATAAGTTTGTGATTAATTCAGGATGAGCTTCGGTTCGTGTGCAC	388
QY	41	ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp	60
DB	389	ACAGTACCTGATGGTCATGTTTGGCGGTGAGGACTAAGGAAAGCTGACAA	448
QY	61	PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu	80
DB	449	TTTCAAGATGGTGGCAGAGATTTGTTGACCGTTACTCCTTCGATTCGATTCCTTTG	508
QY	81	IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer	100
DB	509	ATTTTAGATATGAAGAAACTCTGCCCTTCAGCGTCTACATTTTCAATTTATCCACTCT	568
QY	101	GlulLeuAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla	120

Db 569 GAGATCAATACCATTCACCGGTCTCATGGATTCGGCTCACCAACCACTTCAAAACGGCC 628
Qy 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerValTyr 140
Db 629 CGTTGTTGTCAGACCTTGAAGATGAAGATGCCGAGGTCACTTCTCTCTCTCTGTAC 688
Qy 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
Db 689 CCATCACCACTTCTCGATCTACAGTACCAAGCAACAAAGGGTATGCTAGTCTCAGCCATC 748
Qy 161 GlnThrLeuPheThrGlyProValLysAlaGluProThrProThrProLysIlePro 180
Db 749 CAACCTTGTCTACTGACCACTTAAGCTGAAGGCCAACGCCCAACCCCAAAATACCT 808
Qy 181 LysLeuArgGlyArgLysLysAsnAlaAspProGluLysIleAsnSerSerAlaPro 200
Db 809 AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 868
Qy 201 ArgAspAspAspProGluLeuArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220
Db 869 CGAGATGATGATCCAGAGAACCGTTCAAGTTCTACGAGAGTCTTCTGCGAAGAGAGA 928
Qy 221 ThrValThrAlaGluGluArgAlaIleAsnAlaLysThrPheGluProThr 240
Db 929 ACCGTGACTGCAGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 988
Qy 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260
Db 989 AACCTTCTTCAGAGTGGTCTGCGACCACTCTATCTATCAGAGGTTGCATCATGTAT 1048
Qy 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyLysSerGlyPheIleLysValGln 280
Db 1049 CTTCTCTCTGGGTTGCTGAGAGGTACCTAAGTGGGATCTCCGGGTTCTCAAGTCCAG 1108
Qy 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300
Db 1109 CTTGCGGAGAAACAAATGGCTGTTCGATGCTCTACAAAGCCGGGAGAGCCAAATTCAGT 1168
Qy 301 GlnGlyTyrTyrGluPheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPhe 320
Db 1169 CAAGATGTGTACGANTTCACTCTAGAGAACCACTTAGGAGAGGAGGAGGAGGAGGAGT 1228
Qy 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340
Db 1229 GAGCTGCTCAGAACAGAGATTCGTTTGTAAAGTCAGAGCCTTTCGAGTCAACGAGTAC 1288
Qy 341 Val 341
Db 1289 GTC 1291

RESULT 3
LOCUS AF289052
DEFINITION Arabidopsis thaliana reduced vernalization response 1 (VRN1) mRNA,
VRN1-Ler allele, complete cds.
ACCESSION AF289052
VERSION AF289052.1 GI:21734795
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
1 (bases 1 to 1538)
REFERENCE
AUTHORS Levy,Y.Y., Messagne,S., Wylne,J.S., Gendall,A.R. and Dean,C.
TITLE Multiple roles of Arabidopsis VRN1 in vernalization and flowering
time control
JOURNAL Science 297 (5579), 243-246 (2002)
MEDLINE 22111275
PUBMED 12114624
REFERENCE
AUTHORS Levy,Y.Y., Gendall,A.R. and Dean,C.

VRN1, a gene required for response to vernalization
Unpublished
3 (bases 1 to 1538)
Levy,Y.Y. and Dean,C.
Direct Submission
Submitted (23-JUN-2000) Molecular Genetics, John Innes Centre,
Colney Lane, Norwich, Norfolk NR4 7UH, UK
JOURNAL
FEATURES
source
1. .1538
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/map="between mi339 and m560B2"
/ecotype="Landsberg erecta"
1. .1538
/gene="VRN1"
/allele="Ler"
312. .1337
/gene="VRN1"
/allele="Ler"
/note="contains two Vp1/ABI3-like B3 domains"
/codon_start=1
/product="reduced vernalization response 1"
/protein_id="AAW76973.1"
/db_xref="GI:21734796"
/translation="MPPFFHKLIPSTIQEKRLRVPDKFVSKFDELSVAVALTVPD
GHVWVGRLKADNKIKFQDQWQEFVDRYSIRIGYLLIFRYGNSAPSVYIFNLHSEI
NYHSTGLMDSAHNFKRARLPEDLEDAEVIWPSVYVPSLPSTVPANKYASAI
QTLTGPVKABEPTPTPKIPKRGKKNADPEEINSAPRDDDPENRSKFYESASAI
KRTVTAERERAINAAKTPEPTNPFVLRPSYLRGIMYLPSPGAPKYLUSGSGF
IKVQLAERQWPFVRCLYKAGRAKFSQWYEFTLENNLGGEDVCFELLARTRDPVLKVT
FRVNEYV"
ORIGIN
Alignment Scores:
Pred. No.: 2,83e-139 Length: 1538
Score: 1796.00 Matches: 341
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-10-088-187A-11 (1-341) x AF289052 (1-1538)
Qy 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20
Db 312 ATGCCACGCCCTTCTTCATAAGTTGATTTTCTCATCCACTATCCAGAAAAACGTCGT 371
Qy 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
Db 372 AGGGTCCAGATAGTTTGTGATTAATTCAGGATGAGCTTTCGGTGTCTGTGACATC 431
Qy 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
Db 432 ACAGTACCTGATGGTCATGTTTGGCGTGTAGGACTAAGCAAGCTGACAAACAAATTTGG 491
Qy 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrIleLeu 80
Db 492 TTTCAAGATGGTGGCAAGAGTTTGTTCACCGTTACTCCATTCGATTTGTTATCTTTG 551
Qy 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
Db 552 ATTTTATGATATGAAGAACTCTCGCTTACGGCTCTACATTTTCAATTTATCCACTCT 611
Qy 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
Db 612 GAGATCAATTACCATTCACCGGTCTCATGATTCGCTCAACCACTTCAACACGCCGCC 671
Qy 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140
Db 672 CGTTTGTGAGACCTTTGAAGATGAAGATGCCGAGGTCACTCTTCTCTCTCTCTGTAC 731

```
QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
DB 732 CCATCACCACCTTCTCTGAGTCTACAGTACCAGCCAAACAAAGGATATGCTAGTTCAGCCATC 791
QY 161 GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro 180
DB 792 CAAACCTTGTTCACTGGACCACTTAAAGCTGAAGAGCCAAACGCCAACCCCAAAATACCT 851
QY 181 LysLysArgGlyArgLysLysLysAlaAspProGluGluIleAsnSerSerAlaPro 200
DB 852 AAAAAAGAGAGGAGGAGGAAGAAATAATGCTGATCCTGAGGAATAAATCACTCAGCTCCG 911
QY 201 ArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220
DB 912 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTCTTCTCGAGAAAGAGA 971
QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240
DB 972 ACGTGACTGCAGAGAAAGAGAGAGAGCCATCAATGCAGCCAAACGTTGCAACCAACA 1031
QY 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260
DB 1032 AACCCCTTCTTCAGAGTGGTCTCGCAGCCATCTATCTATACAGAGGTTCATCATGTAT 1091
QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280
DB 1092 CTTCCCTCTGGGTTGCTGAGAGTACCTTAAGTGGAGTCTCGGGTTCTATCAAAGTCCAG 1151
QY 281 LeuAlaGluLysGlnTrpProValArgCysLysLeuTyrLysAlaGlyArgAlaLysPheSer 300
DB 1152 CTTGGGAGAAACAATGGCTCTGTCATCTCTACAAAGCGGAGAGCCAAATTCAGT 1211
QY 301 GlnGlyTyrTyrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPhe 320
DB 1212 CAAGGATGTACGAATTCCTCTAGAGAAACAATAGGAGAGGAGAGCGTCTGTGTGTT 1271
QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340
DB 1272 GAGCTGCTCAGAACCCAGAGATTTCGTTTGAAGTGACAGCCCTTCGAGTCAACGAGTAC 1331
QY 341 Val 341
DB 1332 GTC 1334
```

```
RESULT 4
AX101038 LOCUS 1495 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 12 from Patent WO0121822.
ACCESSION AX101038
VERSION AX101038.1 GI:13619894
```

```
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
```

```
REFERENCE
1 Dean,C. and Levy,Y.Y.
METHODS Methods and means for modification of plant flowering
```

```
JOURNAL
Patent: WO 0121822-A 12 29-MAR-2001;
Plant Bioscience Limited (GB)
```

```
FEATURES
Source
1. .1495
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="vrnl-1 mutation"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 4,77e-138 Length: 1495
Score: 1781.00 Matches: 340
Percent Similarity: 99.71% Conservative: 0
Best Local Similarity: 99.71% Mismatches: 1
```

```
Query Match: 99.16% Indels: 0
DB: 6 Gaps: 0
US-10-088-187a-11 (1-341) x AX101038 (1-1495)
QY 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20
DB 269 ATGCCACGCCCTTCTTCCATAGTTGATTTCTCATCCACTATCCAAAGAAACGCTG 328
QY 21 ArgValProAspLysPheValSerLysPheLysAspClnLeuSerValAlaValAlaLeu 40
DB 329 AGGTCGCCAGATAAGTTTGTGAGTAAATTCAGGATGAGCTTTCGGTGTCTGTTCACATC 388
QY 41 ThrValProAspGlyHisValTyrArgValGlyLeuArgLysAlaAspAsnLysIleTyr 60
DB 389 ACAGTACTGATGTGTCAATGTTGACGTGAGGAAAGCTGAGGAAAGCTGACACAAATTTGG 448
QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80
DB 449 TTTCAAGATGGTTGGCAAGAGTTTGTTCAGCGTTACTCCATTCGATTTGGTATCTCTTTTG 508
QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
DB 509 ATTTTATGATATGAAGGAAACTCTGCCTTCAGCGTCTACATTTTCAATTTATCCACTCT 568
QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
DB 569 GAGATCAATTACCATTCACCGGTCTCATGAGTTCCGTCACCAACCACTTCAACCGGCC 628
QY 121 ArgLeuPheGluAspLeuGluAspAlaGluValIlePheProSerSerValTyr 140
DB 629 CGTTTGTGTAAGACCTTGAAGATGAAGATGCCGAGGTTCATCTTCTCTCTGTGTAC 688
QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
DB 689 CCATCACCATCTCTGAGTCTACAGTACCAAGCCAAAGAGGATGCTAGTTTCAGCCATC 748
QY 161 GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro 180
DB 749 CAAACCTTGTTCATCGGACCACTTAAAGCTGAAGAGCCACGCCCAACCCCAAAATACCT 808
QY 181 LysLysArgGlyArgLysLysLysAlaAspProGluGluIleAsnSerSerAlaPro 200
DB 809 AAAAAGAGAGGAGGAGAGAAAGAAATGCTGATCCTCAGGAAATAAACTCATCAGCTCCG 868
QY 201 ArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220
DB 869 CGAGATGATATCCAGAGAACCGTTCAAAGTTCTACGAGAGTCTTCTCGAGAAAGAGA 928
QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240
DB 929 ACCGTGACTGCAGAAAGAAAGAGAGAGCCATCAATGCAGCCAAACGTTCCAAACCAACA 988
QY 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260
DB 989 AACCCCTTCTTCAGAGTGGTTCGCGACCACTTCATCTATACAGAGGTTCGATCATGTAT 1048
QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280
DB 1049 CTTCCCTCTGGGTTGCTGAGAGTACCTTAAGTGGATCTCCGGGTTTCATCAAGTCCAG 1108
QY 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300
DB 1109 CTTGGGAGAAACAATGGCTGTTCGATGCTCTACAAAGCGGAGAGCCAAATTCAGT 1168
QY 301 GlnGlyTyrTyrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPhe 320
DB 1169 CAAGGATGGTACGAATTCACCTCTAGAGAAACAATAGGAGAGGAGAGCTCTGTGTGTT 1228
QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340
DB 1229 GAGCTGCTCAGAACCCAGAGATTTCGTTTGAAGTGACAGCCCTTCGAGTCAACGAGTAC 1288
```



```

QY      341 Val 341
Db      1289 GTC 1291

RESULT 5
AX101039
LOCUS   AX101039
DEFINITION Sequence 13 from Patent WO0121822.
ACCESSION AX101039
VERSION  AX101039.1 GI:13619895
KEYWORDS
SOURCE  synthetic construct
ORGANISM synthetic construct
        artificial sequences.
REFERENCE
AUTHORS Dean,C. and Levy,Y.Y.
TITLE    Methods and means for modification of plant flowering
        characteristics
JOURNAL  Patent: WO 0121822-A 13 29-MAR-2001;
        Plant Bioscience Limited (GB)
FEATURES
source
1. .1494
   /organism="synthetic construct"
   /mol_type="unassigned DNA"
   /db_xref="taxon:32630"
   /note="vrn1-2 mutation"

ORIGIN
Alignment Scores:
Pred. No.:      5,68e-137      Length:      1494
Score:          1768.00        Matches:     340
Percent Similarity: 99.71%    Conservative: 0
Best Local Similarity: 99.71%  Mismatches:   1
Query Match:     98.44%       Indels:      1
DB:              6           Gaps:         0

US-10-088-187A-11 (1-341) x AX101039 (1-1494)

QY      1 MetProArgProPheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20
Db      269 ATGCCACGGCCCTTCTTCATTAAGTGAATTTCTATCCACTATCCAGAAAACGTCG 328

QY      21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
Db      329 AGGCTCCACAGATAAGTTGTGAGTAAATCAAGGATGAGCTTCGGTCTGTCGACATC 388

QY      41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
Db      389 ACAGTACCTCGATGGTGCATGTTGGCGGTGAGGACTAAGGAAAGCTGACAAACAAATTTGG 448

QY      61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80
Db      449 TTTCAGATGGTTGGCAGAGTTTGTGACCGTTACTCCATTCGCATTCGATTCGTTATCTTTG 508

QY      81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
Db      509 ATTTTAGATATAGAGAACTCTGCCTTCAGCGTCTACATTTTCAATTTATCCACATCT 568

QY      101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
Db      569 GAGATCAATATACATTTCCACCGGTCTCATGGATTCGCTTCACACCACTTCAAACGCGCC 628

QY      121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140
Db      629 CGTTTGTGTTGAAGACCTTGAAGATGAAGATGCGGAGGTCACTTCCTCTTCCTGTCGTAC 688

QY      141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
Db      689 CCATCACCACCTTCTCGAGTCTACAGTACCCAGCAACAAAGGTATGCTAGTTCAGCCATC 748

QY      161 GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro 180
Db      749 CAAACCTTGTCTACTGGACCACTTAAAGCTGAAGACCAACGCCCAACCCCAAAAATACCT 808

```

```

QY      181 LysLysArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro 200
Db      809 AAAAGAGAGAGGAGGAGGAAGAA-AATGCTGATCTCTGAGGAATAAATCACTCATCAGCTCG 867

QY      201 ArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220
Db      868 CGAGATGATGATCCAGAGAACGTTCAAGATTCTACGAGAGTGTCTTCGCGAGAAAGAGA 927

QY      221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240
Db      928 ACCGTGACTGTCAGAAAGAAAGAGAGAGCCATCAATGCAGCCCAAAACGTTTCGAACCAACA 987

QY      241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260
Db      988 AACCCCTTCTTCAGAGTGGTCTCGGACCATCTCTATATACAGAGTTGCATCATGTAT 1047

QY      261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280
Db      1048 CTTCCCTTCTGGGTTTGTCTGAGAGTACCTAAGTGGGATCTCCGGGTTCATCAAGTCCAG 1107

QY      281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300
Db      1108 CTTGCGGAGAAACAATGGCTGTTCGATGCTCTCAAAAGCCGGAGAGACCAATTCAGT 1167

QY      301 GlnGlyTyrTrpTyrGluPheThrLeuGluAsnLeuGlyGluGlyValCysValPhe 320
Db      1168 CAAGGATGGTAGCAATTCACCTCTAGAACACACTTAGGAGAGAGAGACCTCTGTGTGTTT 1227

QY      321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340
Db      1228 GAGCTGCTCAGAACCCAGAGATTTCGTTTGAAGTGACAGCCCTTCGAGTCAACGAGTAC 1287

QY      341 Val 341
Db      1288 GTC 1290

RESULT 6
AF289051
LOCUS   Arabidopsis thaliana reduced vernalization response 1 (VRN1) gene,
DEFINITION VRN1-Ler allele, complete cds.
ACCESSION AF289051
VERSION   AF289051.1 GI:21734793
KEYWORDS
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Levy,Y.Y., Menesse,S., Mylne,J.S., Gendall,A.R. and Dean,C.
TITLE    Multiple roles of Arabidopsis VRN1 in vernalization and flowering
         time control
JOURNAL  Science 297 (5579), 243-246 (2002)
MEDLINE  22111275
PUBMED   12114624
REFERENCE
AUTHORS Levy,Y.Y., Gendall,A.R. and Dean,C.
TITLE    VRN1, a gene required for response to vernalization
JOURNAL  Unpublished
REFERENCE
AUTHORS Levy,Y.Y. and Dean,C.
TITLE    Direct submission
JOURNAL  Submitted (23-JUL-2000) Molecular Genetics, John Innes Centre,
         Colney Lane, Norwich, Norfolk NR4 7UH, UK
FEATURES
source
1. .5000
   /organism="Arabidopsis thaliana"
   /mol_type="genomic DNA"
   /db_xref="taxon:3702"
   /chromosome="3"
   /map="between mi339 and m560B2"

```

gene	/ecotype="Landsberg erecta" <1879. .>4370 /gene="VRN1" /allele="Ler"
mRNA	Join(<1879. .1938,3120. .3567,3664. .3732,3832. .4031, 4122. .4370) /gene="VRN1" /allele="Ler"
CDS	Join(1879. .1938,3120. .3567,3664. .3732,3832. .4031, 4122. .4370) /gene="VRN1" /allele="Ler" /note="contains two VP1/ABI3-like B3 domains" /codon_start=1 /product="reduced vernalization response 1" /protein_id="AA067972.1" /db_xref="GI:21734794" /translation="MPPFFHKLIFSSTIQEKRLRVPDKFVSKFDELSSVAALTVPD GHVWVGLRKADNKTWFOGWQEFVDVRSIRIGYLLIPRYGNSAFSVIIFNLGSHSEI NYHSTGLMDSANHHFKARLPEDLEDAEVLPSSVYPSPLPSTVPANKGYASAR QTLTGPVKAEPTPTPKIPKGRKKKNADPEEINSAPRDDDDPENRSKFYESAR KRTVTAERERAINAAKTFEPNPFVRLPSYLYRGICIMLPSPFAKLYLSIGSF IKVQLAEQWQMPVRCLYKAGRAKFSQGWYEFTLENNLGEQVDFVLELRTDFVLKVA PRVNEV"
ORIGIN	
Alignment Scores:	
Pred. No.:	1.31e-123 Length: 5000
Percent:	1614.50 Matches: 321
Percent Similarity:	77.16% Conservative: 0
Best Local Similarity:	77.16% Mismatches: 0
Query Match:	89.89% Indels: 95
DB:	8 Gaps: 3
US-10-088-187A-11 (1-341) x AF289051 (1-5000)	
QY	21 AtgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
Db	3120 AGGGTCCGAGTAAGTTGTGAGTAATTCAGGATGACCTTCGGTGTCTGTGACCTC 3179
QY	41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
Db	3180 ACAGTACCTGATGGTCATGTTGGCGGTGAGGACTAAGGAAAGCTGACACAAATTTGG 3239
QY	61 PheGlnAspGlyTrpGlnGluPheValAspArgTrpSerIleArgIleGlyTrpLeuLeu 80
Db	3240 TTTCAAGATGGTTGGCAAGAGTTTGTGACCGGTACTCCATCGCATGGTTATCTTTTG 3299
QY	81 IlePheArgTrpGluGlyAsnSerAlaPheSerValTrpIlePheAsnLeuSerHisSer 100
Db	3300 ATTTTATAGATAGAGGAACTCTGCCCTTCAGGCTGTACATTTTCAATTTATCCCACTCT 3359
QY	101 GluIleAsnTrpHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
Db	3360 GAGATCAATTACCATTCACCGGTCTCATGGATTCCGCTCACAAACCACTTCAACGCGCC 3419
QY	121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerValTrp 140
Db	3420 CGTTTGTGTGAAGACCTTGAAGATGAAGATGCCGAGGTCACTTCTCTCTCTGTGTAC 3479
QY	141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTrpAlaSerSerAlaIle 160
Db	3480 CCATCACCCTTCTCTGAGTCTACAGTACCAGGCAACAAAGGTATGCTAGTTCAAGCATC 3539
QY	161 GlnThrLeuPheThrGlyProValLys----- 169
Db	3540 CAACCTTGTTCTACTGGACCACTTAAAGGTGATATTATTAACCACTGATTCCTTTATC 3599
QY	169 ----- 169
Db	3600 TATCGCTGATACCGCTCTTATCATTTCTTTGAGTTGATGCTTGATATTTTCTTTATCT 3659
QY	170 ---AlaGluGluProThrProThrProLysIleProLysLysArgGlyArgLysLysLys 188
Db	
QY	189 AsnAlaAspPro----- 192
Db	3720 AATGCTGATCTCGTAAAGCACTTTTCTCTTTGAAATGCTTCAGACTCGTTTTCAGAGGA 3779
QY	193 -----GluGluIle 195
Db	3780 TTCACAGATCTTCTTCATGATACATATATCTTTTGTATTTCTCTTACAGAGGAATA 3839
QY	196 AsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTrpGluSerAla 215
Db	3840 AACTCATCAGCTCCGCGAGATGATGATCCAGAGAACCGTTTCAAGTTCTTACGAGAGTGT 3899
QY	216 SerAlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaLys 235
Db	3900 TCTGCGAGAAAGAAACCGTGACTGCGAAGAAAGAGAGAGAGCCATCAATGAGCCAAA 3959
QY	236 ThrPheGluProThrAsnProPheArgValValLeuArgProSerTrpLeuTrpArg 255
Db	3960 ACGTTCGAACCAACAAACCTTTCTTTCAGAGTGGTTCTGCGACCATCTCTATATACAGA 4019
QY	256 GlyCysIleMet----- 259
Db	4020 GGTTCATCATGTTGTAATAAAAAACATCTTAGAAGACTTAACTTATTCGGTGTCTTCAC 4079
QY	260 -----TyrLeuProSerGlyPhe 265
Db	4080 TGATCTCTAAAGAGCCTTCTGTTCTGTTCTCTCAACAGATATCTTCTTCTGGGTTT 4139
QY	266 AlaGluLysTrpLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGluLysGln 285
Db	4140 GCTGAGAAGTACCTTAAGTGGATCTCCGGGTTTCATCAAGCTCAGCTTGGCGAACA 4199
QY	286 TrpProValArgCysLeuTrpLysAlaGlyArgAlaLysPheSerGlnGlyTrpTrpGlu 305
Db	4200 TGGCTGTGTCATGCTCTCAAGCCGGAGAGCCAAATTCAGTCAAGATGATGACAA 4259
QY	306 PheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArgThr 325
Db	4260 TTCATCTAGAGAACACTTAGGAGAGGAGACGCTCTGTGTGTGTGAGTCTCAGAACC 4319
QY	326 ArgAspPheValLeuLysValThrAlaPheArgValAsnGluTrpVal 341
Db	4320 AGAGATTTGTTTGAAGTGCACAGCCTTTCGAGTCAACGAGTACGTC 4367
RESULT 7	
AP000735/c	79186 bp DNA linear PLN 14-FEB-2004
LOCUS	Arabidopsis thaliana genomic DNA, chromosome 3, TAC clone:K13E13.
DEFINITION	AP000735 BA000014
ACCESSION	AP000735.1 GI:6451856
VERSION	Arabidopsis thaliana (thale cress)
KEYWORDS	Arabidopsis thaliana
SOURCE	Arabidopsis thaliana
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S. Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC DNA Res. 7 (3), 217-221 (2000)
AUTHORS	2 (bases 1 to 79186)
TITLE	Kaneko,T., Katoh,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S. Direct Submission
JOURNAL	Submitted (18-NOV-1999) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp.
MEDLINE	20363099
PUBMED	10907853
AUTHORS	
JOURNAL	

COMMENT

Tel:81-438-52-3935, Fax:81-438-52-3934)
Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd graph.cgi?c=K13E13
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://combio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS Technical University of
Denmark, http://www.cbs.dtu.dk/services/Netgene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://grenlin2ool.iastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is MCB22 and the 3' clone is MHP21.

FEATURES

source

1. .79186
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="K13E13"
/clone_lib="Mitsu TAC"
/ecotype="Columbia"
32. .1558
/note="unnamed protein product; gb|AAC69119.1
gene_id:K13E13.1
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01688.1"
/db_xref="GI:9280309"
/translation="MRTIQEQLPNDLVEILCRVPATSLRLRSTCKAWNRLFQGDRI
LASKHPEKAGLSLRNDYRIFPISNHLGNSPLKSELIDPHSKNSAFAFEI
SRVHEKGLCSOLDSESVVVMPLTGETWIRTGDPKRGSRFDVGYVYQDKRS
WIKSKYLCYRGYKFEIYDFSDSWRLDIIAPRGISIGSELSVSLKNTYFAK
GYTERPTDFLDFYTERKVPVLLPYQSRRLFOASLSVVRDQKLVLLQLDQSS
KTEIWNVIDETTKGASVTKVPLALDPSHLQIGNDGSFPLGDKVVMFCEKLIDE
NKVKDMYIVGDNVTEVGFVDMGCRVILNVPVSLVQIERAGGNRKRGH"
complement (2742. .3977)
/note="unnamed protein product; gb|AA01602.1
gene_id:K13E13.2
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01689.1"
/db_xref="GI:9280310"
/translation="MRFVAPPPRSGNSPSPSGISSEILSRSDPPLPFPSPPLIA
MVVLAAPFLVYIRGLRSLRFLPFRFRWRCKRRLHLSSASASTSSDLRS
FPPFDPSHYSSYPGLDDSVIKTLPFLFYAAACGKPAVGKTSAAACRDCAVCL
LFEFGDYRTLPCLFHAFLECIDELWRSHPCPLCRTLILGAGVLTMPSPFVFLM
APRIPSDLENNAIITRGETTPSKSNWNTIADTINDQEIIRASVEQSPAISRF
ELKRSYFECESSESERVTPMPATVSPWRVRSRWKQSPFGNLSKRVFSRY
RTKSPFFRRSSAGVFPISERIIPATGSSSRRTKSTMPFFRTAPHSNLRCLGDP
EALLSPFRWRRTDTCRAEN"
complement (Join (5713. .5792,5883. .6006,6112. .6349,
6579. .6697,6873. .6931,7161. .7326,7623. .7682))
/note="unnamed protein product; gene_id:K13E13.3
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01690.1"
/db_xref="GI:9280311"
/translation="MEFVVEEGLVNEDCSTLVLPALSTGNVQLAVDLLVSSYTAER
VGYLDPLPCVNDAYGPLPCGIALPLEVYESSSTATLQAQRSPVAKGMITKFA
ENIAFASSGKKHIIWJSSLDFOQLHNDMSRGPQVYLLSNAESDGRDHCERLGF
RLHEDYSEGRCKWYLSVPEKNSVBEALPSEDELEIDYYPSPFAALFAFKARGL

CDS

join (21250. .21381,21635. .21952,22001. .22061,22130. .22200)
/note="unnamed protein product; gb|AAC72857.1
gene_id:K13E13.7
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01692.1"
/db_xref="GI:9280313"
/translation="MVTQNTKARVTSVSHRRSQDDPSPVKKPKFLVLPTMKDKMM
KIPRFVKLQGSLSSEVLTETPAGFKRSIKLRIGERIWFHEGHSSEAEASIEGH
FLPFEYKNSFRVIFPNVSACETKYPLDAVHIIDSDDDIIDITCKEPTDEHKSRS
RDIELDKLHLDVWNPVHLKESEEDKRVFRG"
complement (22331. .23749)
/note="unnamed protein product; gb|AA015933.1
gene_id:K13E13.8
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01693.1"
/db_xref="GI:9280314"
/translation="MSSVFPGRFLSLQONSKTLLQAKIHAQLVINGCHDNSLFGK
LIGHYCKSPSTESSKLAHLVFPFRGHPKFLFNTLLCKSPEDSIRIFANYASKSS
LLYLNERTFVVLGACASASSALRVGRIVHGMVKGLFELYESELIGTLLHFPYAKN
GDLRYARKVDEMPERTSVTNWMLGGYCSHDKGNHARAKVLPFRPSCGSGVRP
TDTTWCVLISAISQTLLEISLGHVHYTEKLGFTPEVDVFIGTALVDMYKCGCLNNA
FSVLEMLKRVNFTMTATGALNGRNETPNLNRMAESGINKNEITFTSLLSAYR
HIGLVEGIEFLKSMKTGFTVPVIEHYGCTVLLGKAGRIQEAQYFTLAMPKEDAI
LLRSLGKMSKTYGVTGCEITGKALLEREDKLSGSECEYVALSVLAHKGKWE
VEKLKEMKERRIKTRPGYSFV"
25146. .26393
/note="unnamed protein product; gb|AAC69119.1
gene_id:K13E13.9
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01694.1"
/db_xref="GI:9280315"
/translation="MTIPDLNDLVEILCRVPARNLKRSLSTKRWNLFPKDRRFA
REHMKAPKEYLPLMLSEYRIPCPSINLQDVPVSVLKRSLSPDPYQFOIDGRV
FHCGLGLCNVHGKPNRYSKI VVWNPVTGTGTWTEAGYRWKEVRFVLGVCYQODE
NNSCSKKYKILCYPNGQDTEIYELNYSRWTRTIPDGLTIPGWTLYSQTWSMG
NLYIFASKEPLHGLVSLRPFSTKSSCLVTLPYQRPRIEYLPUSIAVRGENJLL
LOLDFSEKTEITWNKIDDTTTKGAASVWTKVLPADLSPLQLFSEYVFLDEDDKY
AVCCERMLPEQHEHRYQCRREYKTKIDTKIYILGSDNKNVDEVGSGEGEATDLSLETSQVI

CDS

join (21250. .21381,21635. .21952,22001. .22061,22130. .22200)
/note="unnamed protein product; gb|AAC72857.1
gene_id:K13E13.7
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01692.1"
/db_xref="GI:9280313"
/translation="MVTQNTKARVTSVSHRRSQDDPSPVKKPKFLVLPTMKDKMM
KIPRFVKLQGSLSSEVLTETPAGFKRSIKLRIGERIWFHEGHSSEAEASIEGH
FLPFEYKNSFRVIFPNVSACETKYPLDAVHIIDSDDDIIDITCKEPTDEHKSRS
RDIELDKLHLDVWNPVHLKESEEDKRVFRG"
complement (22331. .23749)
/note="unnamed protein product; gb|AA015933.1
gene_id:K13E13.8
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01693.1"
/db_xref="GI:9280314"
/translation="MSSVFPGRFLSLQONSKTLLQAKIHAQLVINGCHDNSLFGK
LIGHYCKSPSTESSKLAHLVFPFRGHPKFLFNTLLCKSPEDSIRIFANYASKSS
LLYLNERTFVVLGACASASSALRVGRIVHGMVKGLFELYESELIGTLLHFPYAKN
GDLRYARKVDEMPERTSVTNWMLGGYCSHDKGNHARAKVLPFRPSCGSGVRP
TDTTWCVLISAISQTLLEISLGHVHYTEKLGFTPEVDVFIGTALVDMYKCGCLNNA
FSVLEMLKRVNFTMTATGALNGRNETPNLNRMAESGINKNEITFTSLLSAYR
HIGLVEGIEFLKSMKTGFTVPVIEHYGCTVLLGKAGRIQEAQYFTLAMPKEDAI
LLRSLGKMSKTYGVTGCEITGKALLEREDKLSGSECEYVALSVLAHKGKWE
VEKLKEMKERRIKTRPGYSFV"
25146. .26393
/note="unnamed protein product; gb|AAC69119.1
gene_id:K13E13.9
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01694.1"
/db_xref="GI:9280315"
/translation="MTIPDLNDLVEILCRVPARNLKRSLSTKRWNLFPKDRRFA
REHMKAPKEYLPLMLSEYRIPCPSINLQDVPVSVLKRSLSPDPYQFOIDGRV
FHCGLGLCNVHGKPNRYSKI VVWNPVTGTGTWTEAGYRWKEVRFVLGVCYQODE
NNSCSKKYKILCYPNGQDTEIYELNYSRWTRTIPDGLTIPGWTLYSQTWSMG
NLYIFASKEPLHGLVSLRPFSTKSSCLVTLPYQRPRIEYLPUSIAVRGENJLL
LOLDFSEKTEITWNKIDDTTTKGAASVWTKVLPADLSPLQLFSEYVFLDEDDKY
AVCCERMLPEQHEHRYQCRREYKTKIDTKIYILGSDNKNVDEVGSGEGEATDLSLETSQVI

CDS

join (21250. .21381,21635. .21952,22001. .22061,22130. .22200)
/note="unnamed protein product; gb|AAC72857.1
gene_id:K13E13.7
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01692.1"
/db_xref="GI:9280313"
/translation="MVTQNTKARVTSVSHRRSQDDPSPVKKPKFLVLPTMKDKMM
KIPRFVKLQGSLSSEVLTETPAGFKRSIKLRIGERIWFHEGHSSEAEASIEGH
FLPFEYKNSFRVIFPNVSACETKYPLDAVHIIDSDDDIIDITCKEPTDEHKSRS
RDIELDKLHLDVWNPVHLKESEEDKRVFRG"
complement (22331. .23749)
/note="unnamed protein product; gb|AA015933.1
gene_id:K13E13.8
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01693.1"
/db_xref="GI:9280314"
/translation="MSSVFPGRFLSLQONSKTLLQAKIHAQLVINGCHDNSLFGK
LIGHYCKSPSTESSKLAHLVFPFRGHPKFLFNTLLCKSPEDSIRIFANYASKSS
LLYLNERTFVVLGACASASSALRVGRIVHGMVKGLFELYESELIGTLLHFPYAKN
GDLRYARKVDEMPERTSVTNWMLGGYCSHDKGNHARAKVLPFRPSCGSGVRP
TDTTWCVLISAISQTLLEISLGHVHYTEKLGFTPEVDVFIGTALVDMYKCGCLNNA
FSVLEMLKRVNFTMTATGALNGRNETPNLNRMAESGINKNEITFTSLLSAYR
HIGLVEGIEFLKSMKTGFTVPVIEHYGCTVLLGKAGRIQEAQYFTLAMPKEDAI
LLRSLGKMSKTYGVTGCEITGKALLEREDKLSGSECEYVALSVLAHKGKWE
VEKLKEMKERRIKTRPGYSFV"
25146. .26393
/note="unnamed protein product; gb|AAC69119.1
gene_id:K13E13.9
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01694.1"
/db_xref="GI:9280315"
/translation="MTIPDLNDLVEILCRVPARNLKRSLSTKRWNLFPKDRRFA
REHMKAPKEYLPLMLSEYRIPCPSINLQDVPVSVLKRSLSPDPYQFOIDGRV
FHCGLGLCNVHGKPNRYSKI VVWNPVTGTGTWTEAGYRWKEVRFVLGVCYQODE
NNSCSKKYKILCYPNGQDTEIYELNYSRWTRTIPDGLTIPGWTLYSQTWSMG
NLYIFASKEPLHGLVSLRPFSTKSSCLVTLPYQRPRIEYLPUSIAVRGENJLL
LOLDFSEKTEITWNKIDDTTTKGAASVWTKVLPADLSPLQLFSEYVFLDEDDKY
AVCCERMLPEQHEHRYQCRREYKTKIDTKIYILGSDNKNVDEVGSGEGEATDLSLETSQVI

CDS
LNYPALVQIEQAGGKTKRGDD"
complement (join(27709..27957,28048..28247,28347..28409,
28512..28692))
/note="unnamed protein product; gb|AAD43153.1
gene_id:K13E13.10
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01695.1"
/db_xref="GI:9280316"
/translation="MDSAHNHFKARLPEDLEDAEIVRPSSVVPSPLESTVPANK
GYASAIQTLFGPVKSTPTPKPKGGRKKKNADPEEINSSAPRDDDENSKYE
SASARKIVTAERERAINAAKTPEPTFRFVRISYLYRGCMILPSPGFAEKYLS
GTSFKIVQAEKQWPVRLYKAGRAKFSQGWYEFTLENNIGEDVVCVPELLTRDFV
LKVTAPRVNEYV"
complement (join(32300..32554,32669..32996,33144..33619))
/note="gene_id:K13E13.11"
/codon_start=1
/evidence=not experimental
/product="oxylase-like protein"
/protein_id="BAB01696.1"
/db_xref="GI:9280317"
/translation="MGELDEAFIOAEPHRPNTHLNSGDFISDEIPTIDLSLEDTH
HNKTALEKIAEACKWGFQVNHGLPSALHRVETAAEFNLTTEKKVKRDEV
NEMGHDEHITKNVDWEIEFDPLQDSTIVPASPEDEDTSLRLKLTQWPNQPSHFE
VCOEYAREVEKLAFLLELVLISLGLPGRDRTGFNEQTSFLRFNHPPCNPBLALG
VGRHKDGGALTVLAQDSVGLQVSRSDQWIPVKPISDALIINNQCIVQWINDETW
SABHRVNVNTSKERFSIPFPFSPSEANIEPLEELISEENPCYKYNWGWFFVSRNR
SDFKKLEVENIQIDHFKA"
complement (join(34938..35198,35304..35631,36034..36494))

Alignment Scores:
Pred. No.: 3,98e-122 Length: 79186
Score: 1614.50 Matches: 321
Percent Similarity: 77.16% Conservative: 0
Best Local Similarity: 77.16% Mismatches: 0
Query Match: 89.89% Indels: 95
DB: 8 Gaps: 3
US-10-088-187a-11 (1-341) x AF000735 (1-79186)
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
DB 28959 AGGTCCTCCAGATAAGTTTGTGAGTAATTCAGAGATGAGCTTCGGTTCGTGTCACATC 28900
QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
DB 28899 ACAGTACCTGATGTCATGTTTGGCGGTGTAGGACTAAGGAAAGCTGACAAACAAATTTGG 28840
QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTrpSerIleArgIleGlyTrpLeuLeu 80
DB 28839 TTTCAAGATGTTGGCAGAGTGTGTTGACCGTTACTCCATTCGCAATGCTTATCTTTTG 28780
QY 81 IlePheArgTrpGluGlyAsnSerAlaPheSerValTrpIlePheAsnLeuSerHisSer 100
DB 28779 ATTTTATAGATGAAGAACTCTGCCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 28720
QY 101 GluIleAsnTrpHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
DB 28719 GAGATCAATTAACCATTCACCGGCTCATGGATTCGCTCAACCACTTCAACCGCC 28660
QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTrp 140
DB 28659 CGTTTGTGTGAGACCTTGAAGATGAAGATGCGGAGTTCATCTTCTCTCTGTTGATC 28600
QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTrpAlaSerSerAlaIle 160
DB 28599 CCATCACCACCTCTCTGAGTCTACAGTACCAGCCAAACAAAGGATGCTAGTTCCAGCATC 28540
QY 161 GlnThrLeuPheThrGlyProValLys----- 169
DB 28539 CAAACCTTGTTCACTGGACCAAGTTAAAGGTGATATTTATAACCAACTGATTCCTCTTATC 28480

QY 169 ----- 169
DB 28479 TATCGCTGATTAGCGGCTTATCATCTTTTTCAGGTTGATGCTTATCTTCTTATCTTCTTATCT 28420
QY 170 ---AlaGluGluProThrProThrProLysLysPheValProLysLysArgGlyArgLysLysLys 188
DB 28419 CCAGCTGAAGAGCCACGCAACCCCAAAAATACCTTAAAGAGAGAGGAGGAGGAGGAGGAGG 28360
QY 189 AsnAlaAspPro----- 192
DB 28359 AATGCTGATCTCGTAAAGCACTTTTCTCTTTGAAATGCTTCCAGACTCGTTTTCAGAGG 28300
QY 193 -----GluGluIle 195
DB 28299 TTCACAGATCTCTCTCATGATACATATATCTTTTGTATGTTGCTTACAGAGGAAATA 28240
QY 196 AsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTrpGluSerAla 215
DB 28239 AACTCATCAGTCCGCGAGATGATGATCCAGAGAACCGTTTCAAGTTTCTACGAGAGTGT 28180
QY 216 SerAlaArgLysArgThrValThrAlaGluGluArgGluAlaIleAsnAlaLys 235
DB 28179 TCTGCGAAGAAAGAACCGTGACTGCGAAGAAAGAGAGAGAGCCATCATGCGAGCAAA 28120
QY 236 ThrPheGluProThrAsnProphePheArgValValLeuArgProSerTrpLeuTrpArg 255
DB 28119 ACGTTCGAACCAACAAACCTTTTCTTTCAGAGTGGTTCTGCGACCATCTCTATATACAG 28060
QY 256 GlyCysIleMet----- 259
DB 28059 GGTTCATCATGTGTAATAAAAAACATCTTAGGAGACTTAATCTTATCGGTGCTCTCAC 28000
QY 260 -----TyrLeuProSerGlyPhe 265
DB 27999 TGATCTCTAAAGAGCGCTTCTGTTTCTTCTCTCAACAGATATCTTCTCTCTGCGTTT 27940
QY 266 AlaGluLysTrpLeuSerGlyPheIleLysValGlnLeuAlaGluLysGln 285
DB 27939 GCTGAGAAGTACCTTAAGTGGGATCTCGGGTTTCATCAAGATCCAGCTTGGGAGAACAA 27880
QY 286 TrpProValArgCysLeuTrpLysAlaGlyArgAlaLysPheSerGlnGlyTrpTrpGlu 305
DB 27879 TGGCTGTTTCAATCTCTACAAAGCGCGAGAGCCAAATTCAGTCAAGGATGGTACGAA 27820
QY 306 PheThrLeuGluAsnAsnLeuGlyGlyAspValCysValPheGluLeuLeuArgThr 325
DB 27819 TTCACTCTAGAGAACAACTTAGGAGAGAGAGAGCGTCTGTGTGTGAGCTGCTCAGAAC 27760
QY 326 ArgAspPheValLeuLysValThrAlaPheArgValAsnGluTrpVal 341
DB 27759 AGAGATTTGTTTTGAAAGTGACAGCCTTTTCGATGCAACGAGTACGTC 27712

RESULT 8
AX101027 5000 bp DNA linear PAT 10-APR-2001
LOCUS Sequence 1 from Patent WO0121822.
DEFINITION AX101027
ACCESSION AX101027
VERSION AX101027.1 GI:13619883
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Dean,C. and Levy,Y.Y.
TITLE Methods and means for modification of plant flowering
characteristics
JOURNAL Patent: WO 0121822-A 1 29-MAR-2001;
Plant Bioscience Limited (GB)
FEATURES
Location/Qualifiers
1..5000
/organism="Arabidopsis thaliana"

/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Alignment Scores:
Pred. No.: 1.56e-122 Length: 5000
Score: 1601.50 Matches: 320
Percent Similarity: 76.92% Conservative: 0
Best Local Similarity: 76.92% Mismatches: 1
Query Match: 89.17% Indels: 95
DB: 6 Gaps: 3

US-10-088-187A-11 (1-341) x AX101027 (1-5000)

QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
DB 3120 AGGGTCCAGATAAGTTTGTGAGTAAATTCAGAGATGAGCTTCGGTTCGTGTCACATC 3179

QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
DB 3180 ACAGTACCTGATGTCATGTTTGRCGTGTAGGACTAAGGAAAGCTGACACAAATTTGG 3239

QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrIleuLeu 80
DB 3240 TTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTGCGATTGGTTATCTTTG 3299

QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
DB 3300 ATTTTAGATATGAGGAACCTGCTTCAGGCTCTACATTTTCATTTATCCACTCT 3359

QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
DB 3360 GAGATCAATTACCATTCACCGCTCTCATGGATTCCGCTCACAACCATTTCAACGCGCC 3419

QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140
DB 3420 CGTTTTTTTGAAGACCTTGAAGATGAAGATGCGAGGTTCATCTTCTCTTCTGTGTAC 3479

QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
DB 3480 CCATCACCACTTCCTGAGTCTACAGTACCAGCCACAAAGGTATGTAGTTTCAGCCATC 3539

QY 161 GlnThrLeuPheThrGlyProValLys----- 169
DB 3540 CAAACCTTGTCTACTGGACCACTTAAAGGTATTTATAACCAACTGATCCCTTTATC 3599

QY 169 ----- 169

DB 3600 TATCGCTGATTACGGCTCTTATCATCTTTTGGAGTTGATGCTTGATATTTTCTTATCT 3659

QY 170 ---AlaGluGluProThrProThrProLysIleProLysLysArgGlyArgLysLys 188
DB 3660 CCAGCTGAGAGCCACGACCCCAAAATACCTTAAAGAGAGGAGGAGGAGGAGAA 3719

QY 189 AsnAlaAspPro----- 192
DB 3720 AATGCTGATCTGGTAAGCACTTTTCTCTTTGMAATGCTTCAGACTCGTTTTCAGAGGA 3779

QY 193 -----GluGluIle 195
DB 3780 TTCACAGATTCTTCTCATGATACATATATCTTTTGTATTTGCTTTCAGAGGAGAAATA 3839

QY 196 AsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGluSerAla 215
DB 3840 AACTCATAGCTCCGCGAGATGATGATCCAGAGAACCGTTCAAAGTTTTCACAGAGTGTCT 3899

QY 216 SerAlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaLys 235
DB 3900 TCTCGAGAAAGAGAACCGTGACTGCAGAGAAGAGAGAGAGCCATCAATCAGCCAAA 3959

QY 236 ThrPheGluProThrAsnProPhePheArgValValLeuArgProSerTyrIleuTyrArg 255
DB 3960 ACGTTTCGAACCAACCAACCTTTCTTTCAGAGTGGTTCTCGACCATCTCTATCTATACAGA 4019

QY 256 GlyCysIleMet----- 259
DB 4020 GGTTCATCATGGTAATAAAAAACATCTTAGGAAGACTTAATCTTATCGGTGCTTTCAC 4079

QY 260 -----TyrIleuProSerGlyPhe 265
DB 4080 TGATCTCTAAAGAAGCCCTTCTGTTTCTCTCTCAACAGTATCTCTCTTCTGGGTTT 4139

QY 266 AlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGluLysGln 285
DB 4140 GCTGAGAGTACTTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGCTTCGGAGAAACAA 4199

QY 286 TrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTrpTyrGlu 305
DB 4200 TGGCCTGTTGATGTCTCTACAAAGCCGGAGAGCAAAATTCAGTCAAGGATGGTACGAA 4259

QY 306 PheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArgThr 325
DB 4260 TTCACTCTAGAGAACAACTTAGGAGAGGAGACGCTCTGTGTTTGTAGCTGCTCAGAAC 4319

QY 326 ArgAspPheValIleLysValThrAlaPheArgValAsnGluTyrVal 341
DB 4320 AGAGATTTCGTTTTGAAAGTGACGCTTTCGAGTCAACGAGTACGTC 4367

RESULT 9
LOCUS AY517929 990 bp mRNA linear PLN 18-JAN-2004
DEFINITION Brassica rapa cultivar Kwonsim reduced vernalization response 1
ACCESSION AY517929
VERSION AY517929.1 GI:40806798
KEYWORDS
SOURCE Brassica rapa (Brassica campestris)
ORGANISM Brassica rapa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 990)
AUTHORS Kwon,S.-J., Park,B.-S., Kim,S.-Y., Choi,H.-S., Lee,M.-C., Kim,J.-S., Lee,S.-I., Lim,K.-B., Kim,J.-A., Lee,M.-R., Jin,Y.-M., Kim,D. and Kim,H.-I.
TITLE Brassica rapa Kwonsim VRN1 mRNA
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 990)
AUTHORS Kwon,S.-J., Park,B.-S., Kim,S.-Y., Choi,H.-S., Lee,M.-C., Kim,J.-S., Lee,S.-I., Lim,K.-B., Kim,J.-A., Lee,M.-R., Jin,Y.-M., Kim,D. and Kim,H.-I.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2003) Genomics Division, National Institute of Agricultural Biotechnology, Seodun-dong 225, Suwon 441-707, Republic of Korea
FEATURES
source location/Qualifiers
1..990
/organism="Brassica rapa"
/mol_type="mRNA"
/cultivar="Kwonsim"
/db_xref="taxon:3711"
1..990
/note="VRN1"
/codon_start=1
/product="reduced vernalization response 1"
/protein_id="AAR92199.1"
/db_xref="GI:40806799"
/translation="MPREFFFKLIFSSTIQSKRLRVPDKFVSFRKDELSSVALVTPDS GHVVRGLRKADNNKNIWFQDQWQEFVDYRSIRIGYLLIFRYGNSAFSVCYLNLPQS EINTHSTGLMDSAGNNHFKRPLFEDLEDAETLHTAISOFTFTGPVKPEATP TQTSVKPKRGKKKNADHPENSSAPRDDDPESRSKFYSASAKRTVNAERERA VNAKTRPTNPFFRVLRLPSLYRGCTIMVLPSPGAEKVLGSGIFIVOLGEKQWPV RCLYAGRAKFSQSGWYETVENNLEGSDVCPELLTRDFVLKVTAIRVNEYV"

CDS
ORIGIN
Alignment Scores:

Pred. No.:	4,51e-116	Length:	990
Score:	1513.00	Matches:	296
Percent Similarity:	88.73%	Conservative:	13
Best Local Similarity:	85.06%	Mismatches:	13
Query Match:	84.24%	Indels:	26
DB:	8	Gaps:	5
US-10-088-187A-11 (1-341) x AY517929 (1-990)			
QY	1	MetProArgProPheHisLysLeuIlePheSerThrIleGlnGluLysArgLeu	20
DB	1	ATGCCAGCCCTTCTTCACAAAGTGTATTTCTCATCCACTATCCAAAGAAAGCGTCTA	60
QY	21	ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu	40
DB	61	AGAGTCCAGATPAAGTTGTGAGTAGATTCAGAGGAGGAGCTATCGGTGCAGTTCACATC	120
QY	41	ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAsp-----AsnLys	58
DB	121	ACAGTACCCGATGGTTCATGTTGGCGTGTAGGACTAAGGAAAGCTGACAAACAACAAG	180
QY	59	IleTrpPheGluAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyr	78
DB	181	ATTGTGTTTCAAGATGGTTGGCAAGAGTTGTGTGACCGTTACTCATCGCATTTGGTTAC	240
QY	79	LeuLeuIlePheArgTyrGluGlyAenSerAlaPheSerValTyrIlePheAsnLeuSer	98
DB	241	CTTTTGATTTTGTATACAGAGCACTCTGCCTTCAGCGTGTGATTTACACTTACCA	300
QY	99	HisSerGluIleAsnTyrHisSerThrGlyLeuMetAspSerAla-----HisAsnHis	116
DB	301	CAGTCCGAGATTAACTTACCATTCACCGGTCTCATGGACTCTGCATCACAAACACAC	360
QY	117	PheLysArgAlaArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePhePro	136
DB	361	TTCAAAACGTCCTCGTTGTGTGAAGACCTTGAAGATGAAGATGCTGAGACTCTGCAC	420
QY	137	SerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAla	156
DB	421	ACC-----	423
QY	157	SerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluGlu-----ProThr	174
DB	424	GCTTCAGCCATCCAGAGTCTTCTCACTGACCTGTATAACCTGGAAGAGGCAACACAA	483
QY	175	ProThrProLysIleProLysArgGlyArgLysLysLysAlaAsp---ProGlu	193
DB	484	CAAACTCAAAAGTTCTTAAAGAGAGGAGGAGAGAAAGAAATGCTGATCATCTGAG	543
QY	194	GluIleAsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGlu	213
DB	544	GAAGTAACTCATCTGCACCGAGAGATGATGACCCAGAGCGCTTCAAAGTTCTACGAG	603
QY	214	SerAlaSerAlaArgLysArgThrValThrAlaGluGluArgLysAlaIleAsnAla	233
DB	604	AGTGCTTCTCCGAGAAAGAGAACAGTGAATGCTGAAGAGAGAGAGAGGGCTTTAATG	663
QY	234	AlaLysThrPheGluProThrAsnProPheArgValValLeuArgProSerTyrLeu	253
DB	664	GCCAAACGTTTGAGCCCAACAACCTTCTTCAGAGTTGTTCTGAGACCGTCTTATCTA	723
QY	254	TyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIle	273
DB	724	TACAGAGTTGTCATCATGATTTGCTTCTGTTTGTGAGAGTACTTAAGTGGATC	783
QY	274	SerGlyPheIleLysValGlnLeuAlaGluLysGlnTrpProValArgCysLeuTyrLys	293
DB	784	TCGGGATTCATCAAGGTCAGCTAGGGGAGAAACAATGGCTGTGAGATGCTATACAAA	843
QY	294	AlaGlyArgAlaLysPheSerGlnGlyTyrTyrGluPheThrLeuGluAsnLeuGly	313
DB	844	GCAGGAGAGCTAAGTTCACTCAAGTTGTGTATGAGTTCACTGTGAGAGATTAACCTAG	903
QY	314	GluGlyAspValCysValPheGluLeuLeuArgThrArgAspPheValLeuLysValThr	333
DB	904	GAAGTGTATGTCGCGTGTGTGAGCTGCTGAGAACAGAGATTCGTTTGAAGTGACG	963
QY	334	AlaPheArgValAsnGluTyrVal	341
DB	964	GCCTATCGAGTCAACGAGTACGTC	987
RESULT 10			
LOCUS	AY356368	990 bp	mRNA linear
DEFINITION	Brassica rapa cultivar Samjin reduced vernalization response 1	PLN 25-AUG-2003	
ACCESSION	AY356368		
VERSION	AY356368.1	GI:33943516	
KEYWORDS	Brassica rapa		
SOURCE	Brassica rapa		
ORGANISM	Brassica rapa		
REFERENCE	1 (bases 1 to 990)		
AUTHORS	Kwon,S.-J., Park,B.-S., Kim,S.-Y., Choi,H.-S., Lee,M.-C., Kim,J.-S., Lee,S.-I., Lim,K.-B., Kim,J.-A., Hong,K.-Y., Lee,M.-R., Jin,Y.-M., Kim,D. and Kim,H.-I.		
TITLE	Submitted (23-JUL-2003) Genomics Division, National Institute of		
JOURNAL	Agricultural Biotechnology, Seodun-dong 225, Suwon 441-707, Republic of Korea		
FEATURES	Location/Qualifiers		
source	1..990		
	/organism="Brassica rapa"		
	/mol_type="mRNA"		
	/cultivar="Samjin"		
	/db_xref="taxon:3711"		
CDS	1..990		
	/note="VRN1"		
	/codon_start=1		
	/product="reduced vernalization response 1"		
	/protein_id="AA055453.1"		
	/db_xref="GI:33943517"		
	/translation="MRPFPFKLIFSSTIOEKLRVDFKFSKFDLSVAVALTPD GHVWRGRKADNNKIKWFQDQWQFVDRYSIRIGYLLIFRYSGNSAFSVCIVNLPOS EINYHSTGLMDSHNNHFKRPLFDEDEDAETLITASAQSFTGPKVKEATP TQTSKVPKGRKKKNADHPENVSSAPRDDPESRSKFYSASRGTVNAEBREKA VNAKTEPTNPFRVVRPSYLYRGCMILPSGFABEKLSGISGFIKVQGEKQWVP RCLYKAGRAKFSQGWYEFTEVNNLGEDVCVFBLLTRDFVLKVTAYRVNEYV"		
ORIGIN			
Alignment Scores:			
Pred. No.:	9.67e-116	Length:	990
Score:	1509.00	Matches:	296
Percent Similarity:	88.51%	Conservative:	12
Best Local Similarity:	85.06%	Mismatches:	14
Query Match:	84.02%	Indels:	26
DB:	8	Gaps:	5
US-10-088-187A-11 (1-341) x AY356368 (1-990)			
QY	1	MetProArgProPheHisLysLeuIlePheSerThrIleGlnGluLysArgLeu	20
DB	1	ATGCCAGCCCTTCTTCACAAAGTGTATTTCTCATCCACTATCCAAAGAAAGCGTCTG	60
QY	21	ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu	40
DB	61	AGAGTCCAGATPAAGTTGTGAGTAGATTCAGAGGAGGAGCTATCGGTGCAGTTCACATC	120
QY	41	ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAsp-----AsnLys	58
DB	121	ACAGTACCCGATGGTTCATGTTGGCGTGTAGGACTAAGGAAAGCTGACAAACAACAAG	180
QY	59	IleTrpPheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyr	78
DB	181	ATTGTGTTTCAAGATGGTTGGCAAGAGTTGTGTGACCGTTACTCATCGCATTTGGTTAC	240
QY	79	LeuLeuIlePheArgTyrGluGlyAenSerAlaPheSerValTyrIlePheAsnLeuSer	98
DB	241	CTTTTGATTTTGTATACAGAGCACTCTGCCTTCAGCGTGTGATTTACACTTACCA	300
QY	99	HisSerGluIleAsnTyrHisSerThrGlyLeuMetAspSerAla-----HisAsnHis	116
DB	301	CAGTCCGAGATTAACTTACCATTCACCGGTCTCATGGACTCTGCATCACAAACACAC	360
QY	117	PheLysArgAlaArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePhePro	136
DB	361	TTCAAAACGTCCTCGTTGTGTGAAGACCTTGAAGATGAAGATGCTGAGACTCTGCAC	420
QY	137	SerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAla	156
DB	421	ACC-----	423
QY	157	SerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluGlu-----ProThr	174
DB	424	GCTTCAGCCATCCAGAGTCTTCTCACTGACCTGTATAACCTGGAAGAGGCAACACAA	483
QY	175	ProThrProLysIleProLysArgGlyArgLysLysLysAlaAsp---ProGlu	193
DB	484	CAAACTCAAAAGTTCTTAAAGAGAGGAGGAGAGAAAGAAATGCTGATCATCTGAG	543
QY	194	GluIleAsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGlu	213
DB	544	GAAGTAACTCATCTGCACCGAGAGATGATGACCCAGAGCGCTTCAAAGTTCTACGAG	603
QY	214	SerAlaSerAlaArgLysArgThrValThrAlaGluGluArgLysAlaIleAsnAla	233
DB	604	AGTGCTTCTCCGAGAAAGAGAACAGTGAATGCTGAAGAGAGAGAGAGGGCTTTAATG	663
QY	234	AlaLysThrPheGluProThrAsnProPheArgValValLeuArgProSerTyrLeu	253
DB	664	GCCAAACGTTTGAGCCCAACAACCTTCTTCAGAGTTGTTCTGAGACCGTCTTATCTA	723
QY	254	TyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIle	273
DB	724	TACAGAGTTGTCATCATGATTTGCTTCTGTTTGTGAGAGTACTTAAGTGGATC	783
QY	274	SerGlyPheIleLysValGlnLeuAlaGluLysGlnTrpProValArgCysLeuTyrLys	293
DB	784	TCGGGATTCATCAAGGTCAGCTAGGGGAGAAACAATGGCTGTGAGATGCTATACAAA	843
QY	294	AlaGlyArgAlaLysPheSerGlnGlyTyrTyrGluPheThrLeuGluAsnLeuGly	313
DB	844	GCAGGAGAGCTAAGTTCACTCAAGTTGTGTATGAGTTCACTGTGAGAGATTAACCTAG	903

Db 170 AGCTCTTATGCAATGGATTCGACAGATGATTCACAAACGTCGTCGATTTGTA 229

QY 124 uAspLeuGluAspGluAspAlaGluValIlePheProSerValTyProSerProLe 144

Db 230 AGATCTGAACTCAAGATGCTAAGTCTATTCATTCG-----AT 271

QY 144 uProGluSerThrValProAlaAsnLysGlyTyZrAlaSerSerAla---lleGlnThrLe 163

Db 272 CCTCGAATCTACTGAACCACTGAATAAAGGTTATGGCGTTCTACAGACATCCAAAGCTT 331

QY 163 upheThrGlyProValLysAlaGluGluProThrProThrProLysLysLeProLysLysAr 183

Db 332 TTTCAAA---GAATCTAAAGCTGAGAA-----ACGCCAAGGTACTCTAGAGAG 379

QY 183 GGLyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSeralProArgAspAs 203

Db 380 AGGAGGAGAGAGAGAGATCTTAATCCCGAGGAGTAACCTTCAACTCCCGGTGGAGA 439

QY 203 pAspProGluAsnArgSerLysPheTyrgLysSerAlaSerAlaArgLysArgThrValTh 223

Db 440 TGACTCAGAGAACCGCTCAAGTCTTACGAGAGTCTTCTGCTAGAAAGAGAACTGTAAAC 499

QY 223 rAlaGluGluArgGluArgAlaIleAsnAlaLysThrPheGluProThrAsnProPh 243

Db 500 TGCAGAGGAGAGAGAGAGAGCGCTCATGACGACCAACATTCGACCAACATCCTTA 559

QY 243 ePheArgValValLeuArgProSerTyZrLeuTyZrArgGlyCysIleMetTyZrLeuProSe 263

Db 560 CTTTAGAGTTGTTCTCGCAACCATCATCTATACAGAGGTTGCATCATGTACTTCCCATC 619

QY 263 rGlyPheAlaGluLysTyZrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGl 283

Db 620 TGGGTTTGTCTGAGAAATCTAAGTGGGATATCTGTTTCATCAAGCTCCAGCTCGGTGA 679

QY 283 uLysGlnTrpProValArgCysLeuTyZrLysAlaGlyArgAlaLysPheSerGlnGlyTr 303

Db 680 GAAACATGCGCAGTGGGCTCTTACAAGCAGGAGAGCTAAGTTTAGCCAAAGGATG 739

QY 303 pTyZrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPheGluLeuLe 323

Db 740 GTATGAGTTTCACTCGAGAACATATAGCGAAGGAGATGATGTGTTTGGAGTACT 799

QY 323 uArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyZrVal 341

Db 800 CAGAACTCGGATTTGTTCTCGAAGTACCGCTTTCGTGTCATGAGTATGTG 854

RESULT 12

AC007504 125021 bp DNA linear PLN 30-OCT-2002

LOCUS Arabidopsis thaliana chromosome I BAC F13F21 genomic sequence,

DEFINITION complete sequence.

AC007504

VERSION AC007504.3 GI:5430745

KEYWORDS HTG.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 125021)

AUTHORS Federici, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P., Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S., Lueros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V., Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 125021)

AUTHORS Federici, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P., Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S., Lueros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V., Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.

TITLE Direct Submission

JOURNAL Submitted (06-MAY-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

REFERENCE 3 (bases 1 to 125021)

AUTHORS Federici, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P., Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S., Lueros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V., Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.

TITLE Direct Submission

JOURNAL Submitted (08-JUN-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

REFERENCE 4 (bases 1 to 125021)

AUTHORS Federici, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P., Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S., Lueros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V., Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.

TITLE Direct Submission

JOURNAL Submitted (09-JUL-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

COMMENT On Jul 9, 1999 this sequence version replaced gi:5019265. e-mail for correspondence: arab@sequence.stanford.edu

Genes with similarity to proteins in the databases are described as 'putative', '-like', or 'similar to'. Genes that have EST similarity but no significant protein similarity are described as 'unknown proteins'. Genes that are annotated based only on gene prediction software are described as 'hypothetical proteins'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/section/index.html), GENSCAN (Chris Burge, & A. Salamov, Sanger Centre, http://genomic.sanger.ac.uk/), NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/NetPlantGene.html) and eMotif (Nevill-Manning, C.G. Wu, T.D. & Brutlag, D.L., http://motif.stanford.edu/projects.html).

FEATURES

source

1..125021

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

/chromosome="I"

/map="mi208"

/clone="F13F21"

/complement(4079..5020)

/gene="F13F21.1"

/complement(join(4079..4465,4631..5020))

/gene="F13F21.1"

/note="Hypothetical Protein"

/codon_start=1

/protein_id="AAD43146.1"

/db_xref="GI:5430745"

/translation="MTNLISGSSYPTANLTFPMQWKIECWLRHAFSPVDETCOMVEI

MTLKFYWEYSIDILAAVDFRLKLCLECYFTDLRLTSKSLRHVRSKIYKLF

KAYKRPSSITSSSQVETLEEDIPAGYGFYFVSQVSGSGKSELDIYGEPTDMA

APRFNVLAYWKNDCRCFKLSMACDVLSPITTVASESFSIGSGVLKYRSLLP

ENICALLCTRNLRGFPKGEEREEVEKEEKEEKEA"

6527..10142

/gene="F13F21.2"

/join(6527..6667,6772..6859,7159..7311,7597..7757,

7840..8514,8612..8939,9020..9240,9369..9546,9658..10142)

/gene="F13F21.2"

/note="Hypothetical Protein"

/codon_start=1

/protein_id="AAD43147.1"

/db_xref="GI:5430747"

/translation="MSENTKVEKTRVFTGACNRRVNVNNGASGLNSFGAONAVAF

CPKTAQILITLPGHKASVNCVTLPTSKFAKAKLDKQYLLSGSDGIIILWELSTL

NNVKKFPYSRDLSPFAVSLSLIPKCKVCLDSICVDTKAIVTLSLAEPLPQNGRF

regions of the genome
 DNA Res. 10 (1), 27-33 (2003)
 MEDLINE 22579290
 PUBMED 12693552
 REFERENCE 2 (bases 1 to 78589)
 Sato, S.
 Direct Submission
 Submitted (12-DEC-2002) Shusei Sato, Kazusa DNA Research Institute,
 Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
 Chiba 292-0818, Japan (E-mail: ssato@kazusa.or.jp,
 URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935(ex.2337),
 Fax: 81-438-52-3934)

FEATURES

source

Location/Qualifiers
 1. 78589
 /organism="Lotus corniculatus var. japonicus"
 /mol_type="genomic DNA"
 /variety="japonicus"
 /db_xref="taxon:34305"
 /chromosomes="2"
 /clone="LjT03H18"
 /clone_lib="LjT library"
 /note="TAC clone: TW0262-synonym: Lotus japonicus"

ORIGIN

Alignment Scores:

Pred. No.: 8.23e-59 Length: 78589
 Score: 850.00 Matches: 195
 Percent Similarity: 48.79% Conservative: 46
 Best Local Similarity: 39.47% Mismatches: 71
 Query Match: 47.33% Indels: 182
 DB: 8 Gaps: 11

US-10-088-187A-11 (1-341) x AP006145 (1-78589)

QY 22 ValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeuThr 41
 Db 16037 ATTCGGGACAAATTTCTGAGGAAATATGCGGTTCACACTTCAAAATTCCTCCCTCACT 16096
 QY 42 ValProAspGlyHisValTyrArgValGlyLeuArgLysAlaAspAsnLysIleTyrPhe 61
 Db 16097 GTTCCAGATGTTACTGTTGGGTGTAGGGTTAAGAGGCCGACACAGAAATTTGGTTT 16156
 QY 62 GluAspGlyTyrGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 81
 Db 16157 GCTGATGGTGGCAAGAAATTTGTCACCGCTACACCATGGCATGGATGTTAGTA 16216
 QY 82 PheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSerGlu 101
 Db 16217 TTTCATGATGAAGGGAAATTCATCTTTTCATTTTCATATTTTCAATATGAGCACTTCTGAG 16276
 QY 102 IleAsnTyrHisSerThr-----GlyLeuMetAspSerAlaHisAsn 115
 Db 16277 GTGAATATCACTCTGCACTGAGAGCCGTAATGAAGGCCCTGCTTTACGAATATATCAT 16336
 QY 116 HisPheLysArgAlaArgLysPheGluAspLeuGluAspGluAspAlaGluValIlePhe 135
 Db 16337 CAT-----ATTTTTCGAAATGGAAGATACAGACTCTTCGAATTCATG 16381
 QY 136 ProSerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysGlyTyr 155
 Db 16382 GATTCCTCT-----CCTTCAAAACCTTACTCTCTGGTTTCATTG---CAAAACCAAGGTTTTT 16432
 QY 156 AlaSerSer-----AlaIleGlnThrLeu 163
 Db 16433 GTGGATCTGTAGATCAACTGACCTGGGAAGAAATCAATCTCTGACGGCAGAAATTTG 16492
 QY 164 PheThrGlyProValLys-----AlaGluGluProThrProThrPro 177
 Db 16493 TTTAATGGTGGTCTAAACTCAATCGTGTGAAGTGGGAGATGAGTGGAGTCTTTCTCT 16552
 QY 177 ----- 177
 Db 16553 TCAAAGGATGCCAAATTCACAGGGCAATCACTCAACAGAGATATAGGTGTTTCAGTTTAAT 16612

QY 178 -----LysIlePro----- 180
 Db 16613 GCTGTTGAATTTAAAGGTCTACTCAAGAATTTGAAATTCGTTTATCTCTAACAATGAGGAA 16672
 QY 181 -----LysLysArgGlyArgLysLysAsnAlaAspProGluGluLeu----- 195
 Db 16673 GGGGTTAGCAAACTGTGAAAGAGGAGGAAGTCTGACCAAGTAAGTTCTTTTAGCCA 16732
 QY 195 ----- 195
 Db 16733 AGTTTACTTTTATGATCTGTAATCGAAATAATGAATTTTCTCATCAATGTAATTTGCA 16792
 QY 196 -----AsnSerSerAlaProArgAspAspAspProGluAsnArgSer 209
 Db 16793 GAACTGTAGATGGTCAGGAACCTTCTGCTGAAATGAAGAGGAGGACAGAAATGCGCTAT 16852
 QY 210 LysPheTyrGluSerAlaSerAlaArgLysArgThrValThrAlaGluGluArgGluArg 229
 Db 16853 AGATTTTATGAAAGTCTTCTGCTAGAAAGAGAACCGTACTGCGAGGAGGAGAGAGG 16912
 QY 230 AlaIleAsnAlaAlaLysThrPheGluProThrAsnProPhePheArgValValLeuArg 249
 Db 16913 GCATTAATGATGCGAAAGCATTTGAACAGCTAATCTCTTCTGTCAGTTGTCTTCGA 16972
 QY 250 ProSerTyrLeuTyrArgGlyCysIleMet----- 259
 Db 16973 CCTCTCTATTTATATAGGGATGCATATGTTGAGTTGTTGATTCAGTTTGTATATTTAT 17032
 QY 259 ----- 259
 Db 17033 CTTTATTCTTCTCAACATCAGGGTTTTAAATTTGCGATCCACAGAGCAATTTGCGCCGC 17092
 QY 259 ----- 259
 Db 17093 AACATAACGGTTTTCAGTCCCAACGGCATCAAACTGCAATTTGTCGGCGCTTCAATCT 17152
 QY 259 ----- 259
 Db 17153 GCATTTGTTGCGCAATTCGCAATGTGACCGCAACTGCAATTTAAACCCCTGTTCAACATG 17212
 QY 260 -----TyrLeuProSerGlyPheAlaGluLys 268
 Db 17213 TCATCTTCATGCTAACACCTTCTGTTTCCAGTATCTGCTTCTGCTTTCAGAAAAAG 17272
 QY 269 TyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGlu-----LysGlnTyrPro 287
 Db 17273 AATTTGAATGGGGTTTTCGGATTCATCAAACTTCAGTTGCGGAGAGAGAGAGTGGCCA 17332
 QY 288 ValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTyrTyrGluPheThr 307
 Db 17333 GTTCGCTGCTATATAGAGAGGTGAGCAAAAGTTAAGCCAGGGTTGGTTGATTCAC 17392
 QY 308 LeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArgThrArgAsp 327
 Db 17393 TTAGAGAACAAATTTAGGGGAGGGGATGCTGTGTGTTTGTAGCTGCTCAGAACGAGGAT 17452
 QY 328 PheValLeuLysValThrAlaPheArgValAsnGluTyrVal 341
 Db 17453 GTAGTGCTGCAAGTTACTTTTATTCGTTTAACTGAAGAAGTG 17494

RESULT 15

AC137825/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

unordered pieces.

AC137825

AC137825.22

GI:50897368

HTG; HTGS PHASE1; HTGS_DRAFT.

Medicago truncatula (barrel medic)

Eukaryota; Viridiplantae; Magnoliophyta; Eudicotyledons; core eudicots;

Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;

AC137825 243777 bp DNA linear HTG 13-AUG-2004
 Medicago truncatula clone mth2-23j1, WORKING DRAFT SEQUENCE, 17

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

1 (bases 1 to 243777)
Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D. and Roe, B.A.

Medicago truncatula BAC Clone mth2-23j1
Unpublished
2 (bases 1 to 243777)
Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D. and Roe, B.A.

Direct Submission
Submitted (03-DEC-2002) Department of Chemistry and Biochemistry, The University of Oklahoma, 620 Farrington Oval, Room 208, Norman, OK 73019, USA

3 (bases 1 to 243777)
Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D. and Roe, B.A.

Direct Submission
Submitted (13-AUG-2004) Department of Chemistry and Biochemistry, The University of Oklahoma, 620 Farrington Oval, Room 208, Norman, OK 73019, USA

On Aug 3, 2004 this sequence version replaced gi:50540779.

Center: Department of Chemistry and Biochemistry
The University of Oklahoma
Center code:UOKNOR

* NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 3281: contig of 3281 bp in length
* 3282 3381: gap of unknown length
* 3382 5860: contig of 2479 bp in length
* 5861 5960: gap of unknown length
* 5961 8953: contig of 2993 bp in length
* 8954 9053: gap of unknown length
* 9054 13155: contig of 4102 bp in length
* 13156 16225: gap of unknown length
* 16226 16325: contig of 2970 bp in length
* 16326 21291: gap of unknown length
* 21292 26668: contig of 5277 bp in length
* 26669 31867: gap of unknown length
* 31868 38427: contig of 5099 bp in length
* 38428 38527: gap of unknown length
* 38528 43888: contig of 5361 bp in length
* 43889 43988: gap of unknown length
* 43989 52130: contig of 8142 bp in length
* 52131 52230: gap of unknown length
* 52231 59771: contig of 7541 bp in length
* 59772 59871: gap of unknown length
* 59872 68004: contig of 8133 bp in length
* 68005 83623: gap of unknown length
* 83624 83723: gap of unknown length
* 83724 98823: contig of 15100 bp in length
* 98824 98923: gap of unknown length
* 98924 115790: contig of 16866 bp in length
* 115791 243777: contig of 127888 bp in length.

Location/Qualifiers
1. 243777
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"

FEATURES
source

ORIGIN

Alignment Scores: 1e-54 Length: 243777
Pred. No.: 808.00 Matches: 187
Score: 54.04% Conservative: 47
Percent Similarity: 43.19% Mismatches: 83
Best Local Similarity: 44.99% Indels: 118
Query Match: 2 Gaps: 9
DB: 2
US-10-088-187A-11 (1-341) x AC137825 (1-243777)
Qy 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
Db 155495 AGCTTCCAGATGATTTATGAGGAATATATGGTGGATATTTCTCAACTGTACTCTC 155436
Qy 41 ThrValProAspGlyHisValTrpArgValGlyLysAlaLysAlaLysLysLysLys 60
Db 155435 ACAGTTCCTGATGCGAGTGTGCGGTGTAATAATGAAAAGGTAGACAAATTTGG 155376
Qy 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeu 80
Db 155375 TTCTTCGCGCTGGAATGAATTTGTTCAAACTATTTCCATAGTACTGGATATCTGT 155316
Qy 81 IlePheArgTyrGluGlyAenSerAlaPheSerValTyrIlePheAenLeuSerHis 100
Db 155315 GTATTCAATATGAGGAAAGTTCGATTCCTGCTGTTAATATTTTAGTCTGCCTCT 155256
Qy 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAenHisPhe---Lys 119
Db 155255 GAGATTAACATCAATCACCCGCACAAAGAAAGTAATGAAGCGCTCTCTCTTTGAAA 155196
Qy 120 AlaArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePhe-----Pro 136
Db 155195 CTTACTATTTTCGAAGAAATGGAAGATGAAGATCTCTTGAATCATGGAATCGTC 155136
Qy 137 SerSerValTyrProSerProLeuProGluSerThrValProAlaAen----- 152
Db 155135 ACGAAGCTTACTCTTAGCTATTGCAAAACAAAGACAGTTTCCGGGTCACTGATA 155076
Qy 153 -----LysGlyTyrAlaSerSerAlaIleGlnThrLysPheThrGly----- 166
Db 155075 ACTCTGCGAAGTCTCGTCCGCTCCAGCACTCAGAAATCTGTTAATGATCTAA 155016
Qy 167 ---ProValLysAlaGluGluProThrProThrPro----- 177
Db 155015 AATAGCATAACTGGGGTGAAGTGGGATACTCTTCCAGAAATGACAAATCGGTAG 154956
Qy 177 ----- 177
Db 154955 AACCAAGTTGACGAGATATAGGTCTCCAGTTTAATGTAGTTGATTTAAAAAG 154896
Qy 178 -----LysIleProLysLysArgGlyArgLys 186
Db 154895 GAAGAGTTGAAGTTGCGCGCGCTACTGATGAAAGGTTAAGAAAACCTGCTGTA 154836
Qy 187 Lys----- 187
Db 154835 AGGAATCAGGCTTAGTTATTTGTTTATACTCGGTTTTCTTTATCTAATCTCG 154776
Qy 188 -----LysAenAlaAapProGluGluIleAen 196
Db 154775 AAATCGAAATTTATTTATCAAAATACTATTTTGAAGT-GCAGATGTCAGAA 154717
Qy 197 SerSerAlaProArgAspAspProGluAenArgSerLysPheTyrGluSerAlaSer 216
Db 154716 TCTGAA-----CATGAAGATGAGTGAATGCGCAATAGGTTTATGAAGTGCAT 154663
Qy 217 AlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAenAlaLysThr 236
Db 154662 GCAAGAAAAGAACTGCAACAGCTGAAGAAAGAGAGAGGCGCCATCAATGCAGCAA 154603


```
QY      237 PheGluProThrAsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGly 256
Db      154602 TTGAACCATCTAATCTTCTCGGGTGTCTACGACCCCTCTTATTTATACAGGGGG 154543
QY      257 CysIleMet----- 259
Db      154542 TGCATAAT -GGTGAATCATGAATCCAACACAAATTTTAAATTTTATTTCTCTATAACAT 154484
QY      260 -----TyrLeuProSerGlyPheAlaGlu 267
Db      154483 TTGTATCTTACATGCTAAATCTTTCTGTTTTCCAGTATCTACCACTTCTGCTTTCAGAG 154424
QY      268 LysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGlu---LysGlnTrp 286
Db      154423 AGAACTTGAATGGCGTTTCAGGAATTAATAACTTCAAATCTCTGATGGGAGACAGTGG 154364
QY      287 ProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTyrTyrGluPhe 306
Db      154363 CCAGTTTCGCTGCTTTTATAGAGTGGTAGAGCTAAGTTAAGCCAGGGGTGGTTTGAGTTT 154304
QY      307 ThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArgThrArg 326
Db      154303 TCGCTGGAGACAAATTTAGGGAGGCGATGTGTGTGTTTGAGCTGGTTGCAACAAG 154244
QY      327 AspPheValLeuLysValThrAlaPheArgValAsnGlu 339
Db      154243 GAGGTAGTGTGTCAGGTACTGTATTTCGTATTACAGAG 154205
```

Search completed: December 30, 2004, 07:58:39
Job time : 5164 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 30, 2004, 06:11:13 ; Search time 570 Seconds

(without alignments)
3140.443 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796

Sequence: 1 MPRPFHKLIFSSTIQEKL.....LLRTRDFVLKVTAFRNVEYV 341

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q/cn2_1/USPTO spool/US1008187/runat_29122004.115842.107/app_query.fasta_1.519
-DB=N_Geneseq_23Sep04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1008187 @CGN 1.1 470 @runat_29122004.115842.107 -NCPUS=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_23Sep04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1796	100.0	1026	12	Adn73922 Thale cre
2	1796	100.0	1495	4	Aaf62446 A thalian
3	1789	99.6	1510	3	Aac39629 Arabidops
4	1783	99.3	1509	3	Aac51591 Arabidops
5	1711.5	95.3	1347	3	Aac51587 Arabidops
6	1601.5	89.2	5000	4	Aaf62416 A thalian

RESULT 1

ADN73922

ID ADN73922 standard; cDNA; 1026 BP.

XX

AC ADN73922;

XX

DT 15-JUL-2004 (first entry)

XX

DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 1817.

XX

KW gene; es; plant; transgenic; E2Fa/Dpa transcription factor;

KW growth regulator; animal feed product; thale cress;

KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

XX

OS Arabidopsis thaliana.

XX

PN WO2004035798-A2.

XX

PD 29-APR-2004.

XX

PF 20-OCT-2003; 2003WO-EP011658.

XX

PR 18-OCT-2002; 2002EP-00079408.

XX

PA (CROP-) CROPDESIGN NV.

XX

PI Inze D, De Veylder L, Vlieghe K;

XX

DR WPI; 2004-348466/32.

DR P-PSDB; ADN73923.

ALIGNMENTS

XX Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
PT more proteins.

XX Claim 1; SEQ ID NO 1817; 134pp; English.

XX This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreduplication, biochemistry, signal
CC transduction, storage lipid mobilisation and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polynucleotide sequence is thale cress CDNA
CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa
CC transcription factor, given in an exemplification of the invention.

XX SQ Sequence 1026 BP; 295 A; 235 C; 228 G; 268 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.57e-176 Length: 1026
Score: 1796.00 Matches: 341
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-10-088-187A-11 (1-341) x ADN73922 (1-1026)

QY 1 MetProArgProPheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20
DB 1 ATGCCAGCCCTTCTTCCATGAAGTTGATTTCTCATCCACTATCCAGAAAAAGCTGTG 60
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
DB 61 AGGGTCCCAAGATAAGTTGTGAGTAAATTCAGAGTAGCTTTCGGTTGCTGTGCTACATC 120
QY 41 ThrValProAspGlyHisValTpaArgValGlyLeuArgLysAlaAspAsnLysIleTpa 60
DB 121 ACAGTACCTGATGCTCATGTTTGGCGGTGAGGACTAAGGAGGAGTGCACAAATAATTTGG 180
QY 61 PheGlnAspGlyTpaGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80
DB 181 TTTCAAGATGTTGGCAAGAGTTTGTTCACCGTTACTCCATTCGCAATGCTTATCTTTTG 240
QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
DB 241 ATTTTATAGATATGAAGAACTCGCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 300
QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
DB 301 GAGATCATTTACATTCACCGTCTCATGGATTCGGTCCACACCACTTCAACCGGCC 360
QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140
DB 361 CGTTTGTTCGAAGACCTTGAAGATGAAGATGCGGAGTGCATCTTCTCTCTCTGTGTAC 420
QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
DB 421 CCATCACCATCTCTCTGAGTCTACAGTACCGAGCAACAAAGGGTATGCTAGTTCAGGCATC 480
QY 161 GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro 180

DB 481 CAAACCTTGTCTACTGCAGCCAGTTAAAGCTGAAGAGCCCAAGCCCAACCAATACCT 540
QY 181 LysLysArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro 200
DB 541 AAAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 201 ArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220
DB 601 CGAGATGATGATCCAGAGAACCGTTTCAAGTTCTACGAGAGTCTTCTCGCAGAGAGAGA 660
QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240
DB 661 ACCGTGACTCAG 720
QY 241 AsnProPhePheArgValValLeuArgProSerTyrIleuTyrArgGlyCysIleMetTyr 260
DB 721 AACCTTTTCTTCAGAGTGGTTCTGCGACCATCTATCTATACAGAGAGTTCATCATGTAT 780
QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyLysSerGlyPheIleLysValGln 280
DB 781 CTTCCTTCTGGTTTCTGAGAGTACCTTAAGTGGGATCTTCGGGTTTCATCAAGTCCAG 840
QY 281 LeuAlaGluLysGlnTpaProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300
DB 841 CTTGCGGAGAGAAACAATGCGCTGTCGATGTCTCTACAAAGCGGAGAGGAGGAGGAGGAG 900
QY 301 GlnGlyTpaTyrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPhe 320
DB 901 CAAGGATGGTACGAATTCATCTCTAGAGAACCACTTAGGAGAGGAGGAGGAGGAGGAGG 960
QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340
DB 961 GAGTGTCTCAGAACCAAGAGATTTCGTTTGAAGTGCAGAGCTTTTCGAGTCAACGAGTAC 1020
QY 341 Val 341
DB 1021 GTC 1023
RESULT 2
AAF62446
ID AAF62446 standard; cDNA; 1495 BP.
XX AAF62446;
XX 05-NOV-2001 (first entry)
XX A thaliana VRN1 coding sequence.
XX VRN1; vernalisation; flowering; crop; ss.
XX Arabidopsis thaliana.
XX Key Location/Qualifiers
CDS 269..1294
FT /*tag= a
FT /product= "VRN1"
XX WO200121822-A1.
XX 29-MAR-2001.
XX 13-SEP-2000; 2000WO-GB003525.
XX 17-SEP-1999; 99GB-00022071.
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX Dean C, Levy YI;
XX WPI; 2001-273467/28.
XX P-PSDB; AAB35491.

PT Novel VRN1 polynucleotide sequence encoding a polypeptide which alters
PT vernalization response of plant in which VRN1 nucleic acid is expressed,
PT useful for influencing and assessing vernalization phenotype of plants.
PS
XX Claim 2; Fig 7; 91pp; English.

CC The present invention provides the protein and coding sequences of
CC Arabidopsis thaliana VRN1. This protein is capable of altering the
CC vernalisation responses of a plant. Also provided are a number of PCR
CC primers used to isolate the sequences. The sequences are useful in the
CC production of crop plants, where they are able to control the timing of
CC flowering, the duration of vernalisation required, the optimum
CC temperature, or even eliminate the need for vernalisation completely. The
CC present sequence is the VRN1 coding sequence

XX
SQ Sequence 1495 BP; 403 A; 322 C; 312 G; 458 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,698-176 Length: 1495
Score: 1796.00 Matches: 341
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-088-187A-11 (1-341) x AAF62446 (1-1495)

QY 1 MetProArgProPheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20
DB 269 ATGCCACGGCCCTTTCTTCATAGTGTGATTTCTCATCCACATATCCAGAAAAGCTGTG 328
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
DB 329 AGGGTCCCAGATAAGTTTGTGAGTAAATCAAGGATGAGCTTTCGGTGTCTGTGCACTC 388
QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
DB 389 ACAGTACCTCGTGGTCATGTTTGGCGGTGAGGACTAAGGAAAGCTGACAAACAAATTTGG 448
QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80
DB 449 TTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTCGGATGTTATCTTTTG 508
QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
DB 509 ATTTTATAGATAGAGAAACTCTGCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 568
QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
DB 569 GAGATCAATTACCAATTCACCGGTCTCATGGATTCGGCTCACACCACTTCAACGCGCC 628
QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140
DB 629 CGTTTGTGTAAGACCTTGAAGATGAAGTCCGAGGTCACTTCCCTCTCTCTGTGTGAC 688
QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
DB 689 CCATCACCACTTCTCGAGTCTACAGTACCAGCCCAACAAAGGTATGCTAGTTTCAGCCATC 748
QY 161 GlnThrLeuPheThrGlyProValLysAlaGluProThrProThrProLysIlePro 180
DB 749 CAACCTTGTCTCACTGGACCACTTAAGCTGAAGACCAACGCCCAACCCCAAAATACT 808
QY 181 LysLysArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro 200
DB 809 AAAAAGAGGGGAGGAGGAAGAAATAATGCTGATCTCTGAGAAATAAATCATCATCAGCTCG 868
QY 201 ArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220
DB 869 CGAGATGATGATCCAGAGAACCGTTCAAGTTCTACGAGGTGCTTCTCGAGAAAGAGA 928
QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240

DB 929 ACCGTGACTGCAGAGAAAGAGAGAGAGCCATCAATGACGAGCAAAACGTTGAAACCAACA 988
QY 241 AenProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260
DB 989 AACCTTTCTTTCAGAGTGGTCTCGGACCATCTATCTATATACAGAGGTTGCAATCATGTAT 1048
QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280
DB 1049 CTTCTTCTGGGTTTGTCTGAGAAGTACCTAAGTGGGATCTCCGGGTTTCATCAAGTCCAG 1108
QY 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgGlyLysPheSer 300
DB 1109 CTTGCGGAGAAACAATGGCTGTTCGATGTCTCTACAAAGCCGGAGAGCCAAATTCAGT 1168
QY 301 GlnGlyTyrTyrGluPheThrLeuGluAsnLeuGluGlyGluGlyAspValCysValPhe 320
DB 1169 CAAGGATGGTACGAATTCCTCTAGAGAACCACTTAGGAGAGGAGACGCTCTGTGTGTTT 1228
QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValLeuGluTyr 340
DB 1229 GAGCTGCTCTCAGAACCAAGAGATTTCTGTTTGAAGTGACAGCCCTTTCGAGTCAACGAGTAC 1288
QY 341 Val 341
DB 1289 GTC 1291
RESULT 3
AAC39629
ID AAC39629 standard; DNA; 1510 BP.
XX
AC AAC39629;
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 25332.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.

Mon Jan 3 13:13:30 2005

us-10-088-187a-11.rng

PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	10-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	16-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	17-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	18-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139463P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151066P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139817P.	PR	30-AUG-1999;	99US-0151303P.
PR	21-JUN-1999;	99US-0139899P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140353P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140354P.	PR	07-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	99US-0140895P.	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	99US-0140823P.	PR	13-SEP-1999;	99US-0153758P.
PR	28-JUN-1999;	99US-0140931P.	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	99US-01411287P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-01411842P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142154P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142055P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142390P.	PR	24-SEP-1999;	99US-0155559P.
PR	09-JUL-1999;	99US-0142803P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-015753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-01443325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	23-JUL-1999;	99US-0145192P.	PR	22-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.


```
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1,466-175 Length: 1510
Score: 1789.00 Matches: 340
Percent Similarity: 99.71% Conservatives: 0
Best Local Similarity: 99.71% Mismatches: 1
Query Match: 99.61% Indels: 0
DB: 3 Gaps: 0

US-10-088-187A-11 (1-341) x AAC39629 (1-1510)

QY 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20
Db 286 ATGCCACGCCCTTCTCCATAAGTTGATTTCTCATCTCACTATCCAGAAAAACGTCG 345
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
Db 346 AGGGTCCCAAGATAAGTTTGTGAGTAAATTCAGAGGATGAGCTTCGGTTGCTGTCACCTC 405
QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
Db 406 ACAGTACCTGATGTCATGTTGGGTGAGGACTAAGAAAGCTGACAAACAAATTTGG 465
QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80
Db 466 TTTCAAGATGTTGGCAAGAGTTTGTGACCGTTACTCCATTCGCATGATTATCTTTG 525
QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
Db 526 ATTTTGTAGATATGAAGAACTCTGCTTTGACGGTCTACATTTTCAATTTATCCACTCT 585
QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
Db 586 GAGATCAATTACATTCACCGGTCTCATGGATTCCGCTCAACACCTTCAACACGCCGC 645
QY 121 ArgLeuPheGluAspLeuGluAspAlaGluValIlePheProSerSerValTyr 140
Db 646 CGTTTGTGTTGAAGACCTTGAAGATGAGATGCGGAGGTATCTTCTCTCTGTGTAC 705
QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
Db 706 CCATCACCACCTTCTGAGTCTACAGTACCAGGCCAACAAAGGTATGCTAGTTTCAGCCATC 765
QY 161 GlnThrLeuPheThrGlyProValLysAlaGluProThrProThrProLysIlePro 180
Db 766 CAATCCTTGTTCACGTGACAGTAAAGCTGAAGAGCCAAACGCCCAACCCCAAAATACCT 825
QY 181 LysLysArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro 200
Db 826 AAAAAGAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 885
QY 201 ArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220
Db 886 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGGTGCTTCTGCGAGAAAGAGA 945
QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240
Db 946 ACCGTGACTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1005
QY 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260
Db 1006 AACCTTTCTTCAGAGTGTCTGCGACCATCTCTATCTATACAGAGGTGTCATCATGTAT 1065
QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280
```

```
Db 1066 CTTCTCTGGGTTTGGCTGAGAAAGTACCTAAGTGGGATCTCCGGGTTTCATCAAGTCCAG 1125
QY 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300
Db 1126 CTTGGGAGAAACAATGGCTGTTCGATGTCTCTACAAAGCCGGGAGAGCCAAATTCAGT 1185
QY 301 GlnGlyTrpTyrGluPheThrLeuGluAsnLeuGluGlyAspValCysValPhe 320
Db 1186 CAAGGATGTTACGAATTCACCTAGAGAACAACTTAGGAGAGAGAGACGCTCTGTGTGTTT 1245
QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValLeuGluTyr 340
Db 1246 GAGCTGCTCAGAACCCAGAGATTTGCTTTTGAAGTGCACAGCCCTTTTCGAGTCAACGAGTAC 1305
QY 341 Val 341
Db 1306 GTC 1308

RESULT 4
AAC51591
ID AAC51591 standard; DNA; 1509 BP.
XX
AC AAC51591;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69076.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
```

Mon Jan 3 13:13:30 2005

us-10-088-187a-11.rng

PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	05-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	06-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139111P.	PR	10-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	12-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149802P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140991P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	07-OCT-1999;	99US-0158232P.
PR	16-JUL-1999;	99US-0144325P.	PR	08-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	12-OCT-1999;	99US-0158233P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145152P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161408P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145911P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.

```
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 6,12e-175 Length: 1509
Score: 1783.00 Matches: 341
Percent Similarity: 99.71% Conservatives: 0
Best Local Similarity: 99.71% Mismatches: 0
Query Match: 99.28% Indels: 1
DB: 3 Gaps: 0

US-10-088-187A-11 (1-341) x AAC51591 (1-1509)

QY 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20
DQ |||||
DB 284 ATGCCACGCCCTTCTTCCATAAGTTGATTTCTCATCCACTATCCAGAAAAAGCTCTG 343
QY 21 -ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAla 40
DQ |||||
DB 344 GAGGGTCCAGATAAGTTTGTGAGTAAATTCAGGATGAGCTTCGGTTGCTGTTCGACT 403
QY 40 uThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTr 60
DQ |||||
DB 404 CACAGTACCTGATGTGTCATGTTGGCGTGTAGACTTAGAGAAAGCTGACACAAATTG 463
QY 60 pPheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeu 80
DQ |||||
DB 464 GTTTCAGATGCTTGGCAAGAGTTTGTTCACCGTTACTCCATTCGATTTGTTATCTTT 523
QY 80 uIlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHis 100
DQ |||||
DB 524 GATTTTATGATATGAAGGAACCTCTGCCTTCAGCGCTACATTTTCAATTTATCCACTC 583
QY 100 rGluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAl 120
DQ |||||
DB 584 TGAGATCAATTAACCATTCACCGGTCTCATGGATTCGCTCACAACCACTTCAAACGGCG 643
QY 120 aArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerVal 140
DQ |||||
DB 644 CGGTTTGTGGAAGACCTTGAAGATGAAGATCCGAGGTCACTTTCTTCTCTGTGTA 703
QY 140 rProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAla 160
DQ |||||
DB 704 CCATCACCACCTCTCTGAGTCTACAGTACACGACCAACAAAGGTATGCTAGTTCAGCCAT 763
QY 160 eGlnThrLeuPheThrGlyProValLysAlaGluProThrProThrProLysIlePr 180
DQ |||||
DB 764 CCACACCTTGTCTACTGGACCAAGTTAAAGCTGAAGAGCCACCAACCCCAAAATACC 823
QY 180 oLysLysArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPr 200
DQ |||||
DB 824 TAAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 883
QY 200 oArgLysAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysAr 220
DQ |||||
DB 884 GCGAGATGATGATCCAGACACCGTTCAAAGTTCTACGAGAGTGCTTCTCGGAGAAAGAG 943
QY 220 gThrValThrAlaGluGluArgGluArgAlaIleAsnAlaLysThrPheGluProTh 240
DQ |||||
DB 944 AACCGTGACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1003
QY 240 rAsnProPheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTy 260
DQ |||||
DB 1004 AAACCTTCTTCAGAGTGTTCTCGGACCATCTATCTATACAGAGGTTGATCATGTA 1063
QY 260 rLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGI 280
DQ |||||
DB 1064 TCTTCTCTCTGCGTTTCTGAGAGTACCTTAAGTGGGATCTCCGGTTCATCAAAGTCCA 1123
QY 280 nLeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSe 300
DQ |||||
DB 1124 GCTTCGGGAGAAACAATGCCCTGTTTCGATGTCTCTACAAAGCCGGAGAGCCAAATTCAG 1183
```

```
QY 300 rGlnGlyTrpTyrGluPheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPh 320
DQ |||||
DB 1184 TCAGGATGGTACGAATTCACCTCTAGAGAACCACTTAGGAGAGGAGACGCTCTGTGT 1243
QY 320 eGluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTy 340
DQ |||||
DB 1244 TGAGCTGCTCAGAACCAAGATTTTCGTTTGAAGTGAAGGCTTCGAGATCAACGAGTA 1303
QY 340 rVal 341
DQ |||||
DB 1304 CGTC 1307

RESULT 5
AAC51587
ID AAC51587 standard; DNA; 1347 BP.
XX
AC AAC51587;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69061.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
```

us-10-088-187a-11.rng

Mon Jan 3 13:13:30 2005

PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147915P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139453P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	15-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139461P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139462P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139463P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140821P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0144803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	13-JUL-1999;	99US-0143542P.	PR	29-SEP-1999;	99US-0156596P.
PR	14-JUL-1999;	99US-0143624P.	PR	04-OCT-1999;	99US-0157117P.
PR	15-JUL-1999;	99US-0144005P.	PR	05-OCT-1999;	99US-0157753P.
PR	16-JUL-1999;	99US-0144085P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144086P.	PR	07-OCT-1999;	99US-0158029P.
PR	19-JUL-1999;	99US-0144325P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144331P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159234P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	18-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160814P.
PR	23-JUL-1999;	99US-0145145P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160981P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			

Alignment Scores:	1.39e-167	Length:	1347
Pred. No.:	1711.50	Matches:	332
Score:	84.34%	Conservative:	2

Db	3180	ACAGTACCTGATGGTCATGTTTGGCGTGTAGAGCTAAGGAAAGCTGACAACTAATTTGG	3239
Qy	61	PheGlnAspGlyTyrGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu	80
Db	3240	TTTCAAGATGGTGGCAAGAGTTTGTTCACCGTTACTCCATTCGCAITGGTTATCTTTTG	3299
Qy	81	IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer	100
Db	3300	ATTTTATAGATAGAGGAACTCTCCCTTCAGCGTCTACATTTTCAATTTATCCCACTCT	3359
Qy	101	GlulIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla	120
Db	3360	GAGATCAATTACCATTCACCGGTCTCATGATTTCCGCTCACCACTTTCAAACGGGCC	3419
Qy	121	ArgLeuPheGluAspLeuGluAspAlaGluValIlePheProSerSerValTyr	140
Db	3420	CGTTTGTGTAAGACCTTGAAGATGAAGATGCCGAGGTCATCTTTCCCTTCTCTGTGTAC	3479
Qy	141	ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle	160
Db	3480	CCATCACCACTTCTCAGTCTACAGTACCAGCCAAACAAAGGATGTAGTTCAGCCATC	3539
Qy	161	GlnThrLeuPheThrGlyProValLys	169
Db	3540	CAAACTTGTCTCAGTGACCACTTAAAGTGATATTATACCAACTGATTCCTTTATC	3599
Qy	169	-----	169
Db	3600	TATCGCTGATTAGCGCTTATCATCTTTTTCAGGTTCATGCTTGATATTTTCCTTATCT	3659
Qy	170	----AlaGluGluProThrProThrProLysIlePheProLysLysArgGlyArgLysLys	188
Db	3660	CCAGCTGTAAGAGCAACGCCAACCCCAAAATACCTTAAAGAGAGAGGAGGAAGAAGA	3719
Qy	189	AsnAlaAspPro-----	192
Db	3720	ATGCTGATCCTGGTAAGCACTTTTCTCTTTGAAATGCTTCAGACTCGTTTTCAGAGGA	3779
Qy	193	-----GluGluIle	195
Db	3780	TTCCACAGATCTTCTCATGATACATATATATCTTTTGATATTGTCTTCACAGAGAAATA	3839
Qy	196	AsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGluSerAla	215
Db	3840	AATCATCAGCTCCGCGAGATGATGATTCAGAGAACCGTTCAAAGTTCTACGAGATGCT	3899
Qy	216	SerAlaArgLysArgThrValThrAlaGluGluArgAlaIleAsnAlaLys	235
Db	3900	TCTGCGAGAAAGAGAACCGTGACTGTCAGAGAAAGAGAGAGAGCCATCAATGCAGCCAA	3959
Qy	236	ThrPheGluProThrAsnProPhePheArgValValIleuArgProSerTyrLeuTyrArg	255
Db	3960	ACGTTCCGAACCAACAAACCCCTTTCTTCAGAGTGGTTCTGCGACCATCTATCATACAGA	4019
Qy	256	GlyCysIleMet-----	259
Db	4020	GGTTGCATCATGGTAATAAAAAACATCTTAGGAGACTTAACTTTATCGGTGCTTCAC	4079
Qy	260	-----TyrIleuProSerGlyPhe	265
Db	4080	TGATCTCTAAAGAGCGTCTGTTTCTGTTTCTCTCAACAGTATCTTCCTCTGGGTTT	4139
Qy	266	AlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGluLysGln	285
Db	4140	GCTGAGAGTACCTTAAGTGGGATCTCCGGTTTATCAAAAGTCCAGTTCGGGAGAACAA	4199
Qy	286	TrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTyrTyrGlu	305
Db	4200	TGGCTGTTCGATGCTCTCAAAAGCGGGAGAGCCAAATTCAGTCAAGGATGGTACGAA	4259
Qy	306	PheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArgThr	325

Db	4260	TTCACTCTAGAGAACAACTTAGGAGAGGAGAGCGTCTGTGTGTTTGAGTGTCTAGAACCC	4319
Qy	326	ArgAspPheValIleuLysValThrAlaPheArgValAsnGluTyrVal	341
Db	4320	AGAGATTTCTGTTTGAAGTAGACGCTTTTCGAGTCAACGAGTACGTC	4367
Db	RESULT 7		
Qy	AAC54934		
Db	ID	AAC54934 standard; DNA; 1394 BP.	
Qy	AC	AAC54934;	
Db	XX		
Qy	18-OCT-2000	(first entry)	
Db	DT		
Qy	DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 79593.	
Db	XX		
Qy	KW	Hybridisation assay; genetic mapping; gene expression control;	
Db	KW	protein identification; signal transduction pathway; metabolic pathway;	
Qy	KW	promoter; termination sequence; ss.	
Db	XX		
Qy	OS	Arabidopsis thaliana.	
Db	XX		
Qy	FN	EP1033405-A2.	
Db	XX		
Qy	PD	06-SEP-2000.	
Db	XX		
Qy	PF	25-FEB-2000; 2000EP-00301439.	
Db	XX		
Qy	PR	25-FEB-1999; 99US-0121825P.	
Db	PR	05-MAR-1999; 99US-0123180P.	
Qy	PR	09-MAR-1999; 99US-0123548P.	
Db	PR	23-MAR-1999; 99US-0125788P.	
Qy	PR	25-MAR-1999; 99US-0126264P.	
Db	PR	29-MAR-1999; 99US-0126785P.	
Qy	PR	01-APR-1999; 99US-0127462P.	
Db	PR	06-APR-1999; 99US-0128234P.	
Qy	PR	08-APR-1999; 99US-0128714P.	
Db	PR	16-APR-1999; 99US-0129845P.	
Qy	PR	19-APR-1999; 99US-0130077P.	
Db	PR	21-APR-1999; 99US-0130449P.	
Qy	PR	23-APR-1999; 99US-0130510P.	
Db	PR	23-APR-1999; 99US-0130891P.	
Qy	PR	28-APR-1999; 99US-0131449P.	
Db	PR	30-APR-1999; 99US-0132048P.	
Qy	PR	30-APR-1999; 99US-0132407P.	
Db	PR	04-MAY-1999; 99US-0132484P.	
Qy	PR	05-MAY-1999; 99US-0132485P.	
Db	PR	06-MAY-1999; 99US-0132486P.	
Qy	PR	06-MAY-1999; 99US-0132487P.	
Db	PR	07-MAY-1999; 99US-0132863P.	
Qy	PR	11-MAY-1999; 99US-0134256P.	
Db	PR	14-MAY-1999; 99US-0134218P.	
Qy	PR	14-MAY-1999; 99US-0134219P.	
Db	PR	14-MAY-1999; 99US-0134221P.	
Qy	PR	14-MAY-1999; 99US-0134370P.	
Db	PR	18-MAY-1999; 99US-0134768P.	
Qy	PR	19-MAY-1999; 99US-0134941P.	
Db	PR	20-MAY-1999; 99US-0135124P.	
Qy	PR	21-MAY-1999; 99US-0135353P.	
Db	PR	24-MAY-1999; 99US-0135629P.	
Qy	PR	25-MAY-1999; 99US-0136021P.	
Db	PR	27-MAY-1999; 99US-0136392P.	
Qy	PR	28-MAY-1999; 99US-0136782P.	
Db	PR	01-JUN-1999; 99US-0137222P.	
Qy	PR	03-JUN-1999; 99US-0137528P.	
Db	PR	04-JUN-1999; 99US-0137502P.	
Qy	PR	07-JUN-1999; 99US-0137724P.	
Db	PR	08-JUN-1999; 99US-0138094P.	
Qy	PR	10-JUN-1999; 99US-0138540P.	
Db	PR	10-JUN-1999; 99US-0138847P.	
Qy	PR	14-JUN-1999; 99US-0139119P.	
Db	PR	16-JUN-1999; 99US-0139452P.	
Qy	PR	16-JUN-1999; 99US-0139453P.	

PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151066P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151080P.
PR 23-JUN-1999;	99US-0140353P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140354P.	PR 31-AUG-1999;	99US-0151438P.
PR 24-JUN-1999;	99US-0140695P.	PR 01-SEP-1999;	99US-0151930P.
PR 28-JUN-1999;	99US-0140823P.	PR 07-SEP-1999;	99US-0152363P.
PR 29-JUN-1999;	99US-0140991P.	PR 10-SEP-1999;	99US-0153070P.
PR 30-JUN-1999;	99US-0141287P.	PR 13-SEP-1999;	99US-0153758P.
PR 01-JUL-1999;	99US-0141842P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0142154P.	PR 16-SEP-1999;	99US-0154039P.
PR 02-JUL-1999;	99US-0142055P.	PR 20-SEP-1999;	99US-0154779P.
PR 06-JUL-1999;	99US-0142390P.	PR 22-SEP-1999;	99US-0155139P.
PR 08-JUL-1999;	99US-0142803P.	PR 23-SEP-1999;	99US-0155486P.
PR 09-JUL-1999;	99US-0142920P.	PR 24-SEP-1999;	99US-0155659P.
PR 12-JUL-1999;	99US-0142977P.	PR 28-SEP-1999;	99US-0156458P.
PR 13-JUL-1999;	99US-0143542P.	PR 29-SEP-1999;	99US-0156596P.
PR 14-JUL-1999;	99US-0143624P.	PR 04-OCT-1999;	99US-0157117P.
PR 15-JUL-1999;	99US-0144005P.	PR 05-OCT-1999;	99US-0157753P.
PR 16-JUL-1999;	99US-0144085P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144086P.	PR 07-OCT-1999;	99US-0158029P.
PR 19-JUL-1999;	99US-0144325P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144331P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0158293P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144334P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159329P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159637P.
PR 21-JUL-1999;	99US-0144814P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0145086P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160741P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160814P.
PR 23-JUL-1999;	99US-0145145P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160981P.
PR 26-JUL-1999;	99US-0145276P.	PR 22-OCT-1999;	99US-0160989P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145919P.	PR 25-OCT-1999;	99US-0161406P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161359P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146388P.	PR 26-OCT-1999;	99US-0161361P.
PR 03-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161920P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161992P.
PR 04-AUG-1999;	99US-0147204P.	PR 28-OCT-1999;	99US-0161993P.
PR 05-AUG-1999;	99US-0147302P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147192P.		
PR 06-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 09-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		

Alignment Scores:
Pred. No.: 2.33e-108
Score: 1141.50
Percent Similarity: 75.93%
Best Local Similarity: 69.63%
Query Match: 63.56%
DB: 3
Length: 1394
Matches: 243
Conservative: 22
Mismatch: 28
Indels: 57
Gaps: 10

US-10-088-187A-11 (1-341) x AAC54934 (1-1394)

Arabidopsis thaliana DNA fragment SEQ ID NO: 18041.
Hybridisation assay; genetic mapping; gene expression control;
protein identification; signal transduction pathway; metabolic pathway;
promoter; termination sequence; ss.

QY	41	ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTyr	60	DE	XX
Db	274	-----CGTCT-GGACTAAGGAAGCCAAACAATAATCTGG	308	KW	XX
QY	61	PheGlnAspGlyTyrGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu	80	KW	XX
Db	309	TTTCAAGACGGTTGGCAGGAGTTTGTCAACGGTTCTCCATTCGGATTGGT-----	359	OS	XX
QY	81	IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeu---Ser-Hi	99	XX	XX
Db	360	-----TTTCAGATACAAA-----GTTACAGTCTACATTTTCATTTATCTCCACA	404	XX	XX
QY	99	sSerGluIleAsnTyrHisSerThrGly-----LeuMetAspSerAlaHisAs	115	XX	XX
Db	405	CTCTGAGATCAACCACTTCTAGTAGTGAAGCTCTTATGCATAATGGATTCGCACAGAA	464	XX	XX
QY	115	nHisPhe---LysArgAlaArgLeuPheGluAspGluAspAlaGluValI1	134	XX	XX
Db	465	TCAGTTTCAACAAACGTCGCTCGATTGTTGAAGATCCTGAACCTCAAGATGCTAAGGTCT	524	XX	XX
QY	134	ePheProSerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysG1	154	XX	XX
Db	525	TTATCCATCGAAC-----CCTGAATCTACTGAACCGATGAATAAAGG	566	XX	XX
QY	154	yTyr---AlaSerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluLupr	173	XX	XX
Db	567	TTATGCGGTCTTACAGCCATCCAAAGCTTTTCAAA---GAATCTAAAGCTGAAGAA--	621	XX	XX
QY	173	othrProThrProLysIleProLysValArgGlyLysLysLysAsnAlaAspProG1	193	XX	XX
Db	622	-----ACGCCAAGGTACTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	674	XX	XX
QY	193	uGluIleAsnSerAlaProArgAspAspProGluAsnArgSerLysPheTyrG1	213	XX	XX
Db	675	GGAGTAAACTCTTCACTCCCGGTGGAGATGACTCAGAGAACCGCTCAAGTCTTACGA	734	XX	XX
QY	213	uSerAlaSerAlaArgLysArgThrValThrAlaGluArgGluArgAlaIleAsnAl	233	XX	XX
Db	735	GAGTGTCTTCTGTAGAAAGAACTGTAACTGCAGAGGAAAGAGAGAGAGAGAGAGAG	794	XX	XX
QY	233	aAlaLysThrPheGluProThrAsnProPheArgValValLeuArgProSerTyrIe	253	XX	XX
Db	795	AGCCAAACATTTGAAACCAACAAATCTTACITTAGAGTTGTTCTGCGACCATCATCT	854	XX	XX
QY	253	uTyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyI1	273	XX	XX
Db	855	ATACAGAGGTTGCATCATGTACTTGGCATCTGGGTTTGTCTGAGAAATACCTAAGTGGAT	914	XX	XX
QY	273	eSerGlyPheIleLysValGlnLeuAlaGluLysGlnTyrProValArgCysLeuTyrIy	293	XX	XX
Db	915	ATCTGGTTTTCATCAAGCTCCAGCTCGGTGAGAAACAAATGGCCAGTGGGTGCTTACAA	974	XX	XX
QY	293	sAlaGlyArgAlaLysPheSerGlnGlyTyrTyrGluPheThrLeuGluAsnLeuG1	313	XX	XX
Db	975	AGCAGGAGAGCTAAGTTTACCAAGATGATGATGATGATGATGATGATGATGATGATG	1034	XX	XX
QY	313	yGluGlyAspValCysValPheGluLeuLeuArgThrArgPheValLeuLysValTh	333	XX	XX
Db	1035	CGAAGGAGATGATGT	1094	XX	XX
QY	333	rAlaPheArgValAsnGluTyrVal	341	XX	XX
Db	1095	CGCTTTCGTGTCAATGATGATGTG	1119	XX	XX
RESULT	9				
ID	AAC37617				
XX	AAC37617 standard; DNA; 1344 BP.				
AC	AAC37617;				
XX					
DT	17-OCT-2000 (first entry)				
XX					

Arabidopsis thaliana.
EP1033405-A2.
06-SEP-2000.
25-FEB-2000; 2000EP-00301439.
25-FEB-1999; 99US-0121825P.
05-MAR-1999; 99US-0123180P.
09-MAR-1999; 99US-0123548P.
23-MAR-1999; 99US-0125788P.
25-MAR-1999; 99US-0126264P.
29-MAR-1999; 99US-0126785P.
01-APR-1999; 99US-0127462P.
08-APR-1999; 99US-0128234P.
08-APR-1999; 99US-0128714P.
16-APR-1999; 99US-0129845P.
19-APR-1999; 99US-0130077P.
21-APR-1999; 99US-0130449P.
23-APR-1999; 99US-0130510P.
23-APR-1999; 99US-0130891P.
28-APR-1999; 99US-0131449P.
30-APR-1999; 99US-0132048P.
30-APR-1999; 99US-0132407P.
04-MAY-1999; 99US-0132484P.
05-MAY-1999; 99US-0132485P.
06-MAY-1999; 99US-0132486P.
06-MAY-1999; 99US-0132487P.
07-MAY-1999; 99US-0132863P.
11-MAY-1999; 99US-0134256P.
14-MAY-1999; 99US-0134218P.
14-MAY-1999; 99US-0134219P.
14-MAY-1999; 99US-0134221P.
14-MAY-1999; 99US-0134370P.
18-MAY-1999; 99US-0134768P.
19-MAY-1999; 99US-0134941P.
19-MAY-1999; 99US-0135124P.
20-MAY-1999; 99US-0135353P.
21-MAY-1999; 99US-0135629P.
24-MAY-1999; 99US-0136021P.
25-MAY-1999; 99US-0136392P.
27-MAY-1999; 99US-0136782P.
28-MAY-1999; 99US-0137222P.
01-JUN-1999; 99US-0137528P.
03-JUN-1999; 99US-0137502P.
04-JUN-1999; 99US-0137724P.
07-JUN-1999; 99US-0138094P.
08-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
10-JUN-1999; 99US-0139119P.
14-JUN-1999; 99US-0139452P.
16-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139457P.
18-JUN-1999; 99US-0139458P.
18-JUN-1999; 99US-0139459P.
18-JUN-1999; 99US-0139460P.
18-JUN-1999; 99US-0139461P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
18-JUN-1999; 99US-0139763P.
21-JUN-1999; 99US-0139817P.

Mon Jan 3 13:13:30 2005

us-10-088-187a-11.rng

Db	256	TTTCAAGACGGTGGCAGGAGTTTGTCAACGGTTTCTCCATTCCGATTGGT-----	306
Qy	81	lIlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeu---Ser-Hi	99
Db	307	---TTCAATACAAA-----GTTACAGTCTACATTTTCATTTATCTCCACA	351
Qy	99	eSerGluIleAsnTyrHisSerThrGly-----LeuMetAspSerAlaHisAs	115
Db	352	CTCTGAGATCAACACCATCTTAGTAGTGAAGCTCTTATGCAATGGATTCCGCACAGAA	411
Qy	115	nHisPhe---lysArgAlaArgLeuPheGluAspLeuGluAspAlaGluValI	134
Db	412	TCAGTTCAACAACGTCGTGATTTTGAAGATCCTGAATCTCAAGATGTAAGTTCAT	471
Qy	134	ePheProSerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysG	154
Db	472	TTATCCATCGAAC-----CCTGAATTTACTGAACGAGTAATAAGG	513
Qy	154	yTyr---AlaSerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluPr	173
Db	514	TTATGCGGTTCTACAGCCATCCAAAGCTTTTTCAAA---GAATCTAAAGCTGAAGAA--	568
Qy	173	oThrProThrProLysIleProLysIleArgGlyArgLysLysAsnAlaAspProG	193
Db	569	-----ACGCCAAGGTACTTAAGAGAGAGAGAGAGAGAGATCCTAATCCCGA	621
Qy	193	uGluIleAsnSerAlaProArgAspAspAspProGluAsnArgSerLysPheTyrG	213
Db	622	GGNAGTAACCTCTTCAACTCCCGTGGAGATGACTCAGAGAACCCTCAAGTTCTACGA	681
Qy	213	uSerAlaSerAlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAl	233
Db	682	GAGTGTCTTCTAGAAAGAGAACTGTAATCGAGAGGAAAGAGAGAGCCGTCATATGC	741
Qy	233	aAlaLysThrPheGluProThrAsnProPhePheArgValValLeuArgProSerTyrIe	253
Db	742	AGCAAAACATTCGACACCAACAAATCCTTACTTTAGAGTTGTTCTCGACCATCATATCT	801
Qy	253	uTyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyI	273
Db	802	ATACAGAGGTTCATCATCTACTTGGCCATCTGGCTTGTCTGAGAAATACCTAAGTGGGAT	861
Qy	273	eSerGlyPheIleIleValGlnLeuAlaGluLysGlnTyrProValArgCysLeuTyrIy	293
Db	862	ATCTGGTTTCATCAAGCTCCAGCTCGGTGAGAAACAAATGCGCAGTGAGTGCTCTACAA	921
Qy	293	sAlaGlyArgAlaLysPheSerGlnGlyTyrTyrGluPheThrLeuGluAsnLeuG	313
Db	922	AGCAGGAGAGCTAAGTTTAGCCCAAGATGGTATGATGATTCACACTCGAGAACAATATAGG	981
Qy	313	yGluGlyAspValCysValPheGluLeuLeuArgThrArgAspPheValLeuLysValTh	333
Db	982	CGAAGGAGATGTATGTGTGTGAGTCTACTCAGAACTCGGGATTCGTTCTCGAAGTCA	1041
Qy	333	xAlaPheArgValAsnGluTyrVal	341
Db	1042	CGCTTTCTGTGTCATGAGTATGTG	1066
RESULT 10			
ID	AAC37903		
XX	AC	AAC37903 standard; DNA; 753 BP.	
XX	AC	AAC37903;	
XX	DT	17-OCT-2000 (first entry)	
XX	DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 19075.	
XX	DE	Hybridisation assay; genetic mapping; gene expression control;	
XX	KW	protein identification; signal transduction pathway; metabolic pathway;	
XX	KW	promoter; termination sequence; ss.	
XX	OS	Arabidopsis thaliana.	

XX	PN	EP1033405-A2.	
XX	PD	06-SEP-2000.	
XX	PF	25-FEB-2000; 2000EP-00301439.	
XX	XX	99US-0121825P.	
PR	PR	99US-0123180P.	
PR	PR	99US-0123548P.	
PR	PR	99US-0125788P.	
PR	PR	99US-0126264P.	
PR	PR	99US-0126785P.	
PR	PR	99US-0127462P.	
PR	PR	99US-0128234P.	
PR	PR	99US-0128714P.	
PR	PR	99US-0129845P.	
PR	PR	99US-0130077P.	
PR	PR	99US-0130449P.	
PR	PR	99US-0130510P.	
PR	PR	99US-0130891P.	
PR	PR	99US-0131449P.	
PR	PR	99US-0132048P.	
PR	PR	99US-0132407P.	
PR	PR	99US-0132484P.	
PR	PR	99US-0132485P.	
PR	PR	99US-0132486P.	
PR	PR	99US-0132487P.	
PR	PR	99US-0132863P.	
PR	PR	99US-0134256P.	
PR	PR	99US-0134218P.	
PR	PR	99US-0134219P.	
PR	PR	99US-0134221P.	
PR	PR	99US-0134370P.	
PR	PR	99US-0134768P.	
PR	PR	99US-0134941P.	
PR	PR	99US-0135124P.	
PR	PR	99US-0135353P.	
PR	PR	99US-0135629P.	
PR	PR	99US-0136021P.	
PR	PR	99US-0136392P.	
PR	PR	99US-0136782P.	
PR	PR	99US-0137222P.	
PR	PR	99US-0137528P.	
PR	PR	99US-0137502P.	
PR	PR	99US-0137724P.	
PR	PR	99US-0138094P.	
PR	PR	99US-0138540P.	
PR	PR	99US-0138847P.	
PR	PR	99US-0139119P.	
PR	PR	99US-0139452P.	
PR	PR	99US-0139453P.	
PR	PR	99US-0139492P.	
PR	PR	99US-0139454P.	
PR	PR	99US-0139455P.	
PR	PR	99US-0139456P.	
PR	PR	99US-0139457P.	
PR	PR	99US-0139458P.	
PR	PR	99US-0139459P.	
PR	PR	99US-0139460P.	
PR	PR	99US-0139461P.	
PR	PR	99US-0139462P.	
PR	PR	99US-0139463P.	
PR	PR	99US-0139750P.	
PR	PR	99US-0139763P.	
PR	PR	99US-0139817P.	
PR	PR	99US-0139899P.	
PR	PR	99US-0140353P.	
PR	PR	99US-0140354P.	
PR	PR	99US-0140695P.	
PR	PR	99US-0140823P.	
PR	PR	99US-0140991P.	
PR	PR	99US-0141287P.	

PD	04-APR-2002.	Score: 512.00	Matches: 96
XX	26-JAN-2001; 2001US-00770423.	Percent Similarity: 97.96%	Conservative: 0
XX	27-JAN-2000; 2000US-0178512P.	Best Local Similarity: 97.96%	Mismatches: 2
XX		Query Match: 28.51%	Indels: 0
XX		DB: 8	Gaps: 0
PA	(GORL/) GORLACH J.	US-10-088-187A-11 (1-341) x ABX62122 (1-426)	
PA	(ANY/) AN Y.		
PA	(HAMI/) HAMILTON C M.		
PA	(PRIC/) PRICE J L.		
PA	(RAIN/) RAINES T M.		
PA	(YUY/) YU Y.		
PA	(RAME/) RAMEAKA J G.		
PA	(PAGE/) PAGE A.		
PA	(MATH/) MATHAW A V.		
PA	(LEDF/) LEDFORD B L.		
PA	(WOES/) WOESSNER J P.		
PA	(HAAS/) HAAS W D.		
PA	(GARC/) GARCIA C A.		
PA	(KRIC/) KRICKER M.		
PA	(SLAT/) SLATER T.		
PA	(DAVI/) DAVIS K R.		
PA	(ALLE/) ALLEN K.		
PA	(HOFF/) HOFFMAN N.		
PA	(HURB/) HURBAN P.		
XX			
PI	Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;		
PI	Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;		
PI	Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;		
PI	Hurban P;		
XX			
DR	WPI; 2003-110411/10.		
XX			
PT	Novel Arabidopsis thaliana nucleic acid useful for identifying homologous		
PT	or related genes, and to create genetically modified and transgenic		
PT	organisms, such as plant cells and plants.		
XX			
PS	Claim 1; SEQ ID NO 237; 43pp; English.		
XX			
CC	The invention describes an Arabidopsis thaliana nucleic acid (I). The		
CC	polypeptide (II) encoded by (I), transgenic plant (III) or genetically		
CC	modified cell (IV) are useful for screening a candidate agent for its		
CC	biological effect, by combining the candidate agent with (II), (III) or		
CC	(IV), and determining the effect of the candidate agent on (II), (III) or		
CC	(IV). (I) is useful for identifying homologous or related genes, for		
CC	producing compositions that modulate the expression or function of its		
CC	encoded protein, for mapping functional regions of the protein, in		
CC	diagnosis, for studying associated physiological pathways, for genetic		
CC	manipulation of cells, preferably plant cells, in screening assays of		
CC	various plant strains to determine the strains that are capable of		
CC	withstanding a particular disease or environmental stress, for enhancing		
CC	or inhibiting production of biosynthetic product in a plant, for		
CC	producing polypeptides, as probes for the detection of mRNA in biological		
CC	samples, to generate additional copies of (I), to generate ribozymes or		
CC	oligonucleotides, as single stranded DNA probes or as triple-strand		
CC	forming oligonucleotides, and to create genetically modified and		
CC	transgenic organisms, such as plant cells and plants. (II) or (III) is		
CC	useful for introducing or improving disease resistance and stress		
CC	tolerance in plants, screening biological active agents, e.g.,		
CC	fungicides, insecticides, etc., and for elucidating biochemical pathways.		
CC	(III) is useful as crops for their enhanced disease resistance, enhanced		
CC	traits of interest, for screening programs, as crops which exhibit		
CC	enhanced tolerance to environmental stress, or to produce a factor. This		
CC	sequence represents a nucleic acid that may correspond to naturally		
CC	occurring Arabidopsis thaliana expressed sequences. Note: The sequence		
CC	data for this patent did not form part of the printed specification, but		
CC	was obtained in electronic format directly from the US patent office at		
CC	seqdata.uspto.gov/sequence.html?DocID=999909770423		
XX			
SQ	Sequence 426 BP; 127 A; 101 C; 82 G; 114 T; 0 U; 2 Other;		
Alignment Scores:			
Pred. No.:	1.03e-43	Length:	426

QY	244 PheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeuProSer	263
DB	425 TTCAGAGTGGTTCGGACCATCTCTATACAGAGGTTCATCATGTATCTCTCTCT	366
QY	264 GlyPheAlaGluTyrLeuSerGlyIleSerGlyPheIleIleValGlnLeuAlaGlu	283
DB	365 GGGTTTGTCTGAGAGTACCTAAGTGGGATCTCCGGGTTTCATCAAAAGTCCAGNATGGAG	306
QY	284 IysGlnTTPProValArgCysLeuTyrIleValAlaGlyArgAlaIysPheSerGlnGlyTTP	303
DB	305 AAACAATGGCCTGTTCGATGTCTCTACAAAGCCGGAGAGCCAAATTCAGTCAAGGATGG	246
QY	304 TyrGluPheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeu	323
DB	245 TACGATTCACCTCTAGAGACCACTTAGGAGAGGAGAGCTCTGTGTGTTGAGCTGCTC	186
QY	324 ArgThrArgAspPheValLeuIleValThrAlaPheArgValAsnGluTyrVal	341
DB	185 AGAACACAGAGATTTCGTTTGAAGTGACAGCGCTTCGAGTCAACGAGTACGTC	132
RESULT 13		
ADC03457		
ID	ADC03457 standard; DNA; 1361 BP.	
XX		
AC	ADC03457;	
XX		
DT	18-DEC-2003 (first entry)	
XX		
DE	Wheat flowering time-related DNA sequence #17.	
XX		
KW	flowering time; wheat; flower architecture; plant biotechnology;	
KX	commercial plant farming; agriculture; flowering-related protein; da.	
XX		
OS	Triticum aestivum.	
XX		
PN	WO2003000904-A2.	
PD	03-JAN-2003.	
XX		
PF	24-JUN-2002; 2002WO-EP006968.	
XX		
PR	22-JUN-2001; 2001US-0300112P.	
PR	26-SEP-2001; 2001US-0325277P.	
PR	30-NOV-2001; 2001US-0334984P.	
XX		
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.	
PI	Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T;	
PI	Glazebrook J, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;	
XX		
DR	WPI; 2003-229340/22.	
XX		
PT	New isolated nucleic acid molecule encoding a polypeptide modulating	
PT	flower architecture and flowering time, useful in the area of plant	
PT	biotechnology, and commercial plant farming and agriculture.	
XX		
PS	Disclosure; SEQ ID NO 97; 323pp; English.	
XX		
CC	The invention comprises the amino acid and coding sequences of proteins	
CC	involved in the control of flowering time in rice. The DNA and protein	
CC	sequences of the invention are useful for modulating flower architecture	
CC	and flowering time, the DNA and protein sequences are useful in the area	
CC	of plant biotechnology, commercial plant farming and agriculture. The	
CC	present wheat DNA sequence shows homology to the rice sequences of the	
CC	invention.	

XX 312 LeuGlyGluGlyAspValCysValPheGluLeuLeuArg 324
SQ 967 TCGCGGAGGGCGAGCTGCTGCTTTTGTGATGATCAAA 1005

Alignment Scores:
Pred. No.: 8,14e-21 Length: 1361
Score: 298.50 Matches: 93
Percent Similarity: 44.71% Conservative: 47
Best Local Similarity: 29.71% Mismatches: 142
Query Match: 16.62% Indels: 31
DB: 10 Gaps: 10

US-10-088-187A-11 (1-341) x ADC03457 (1-1361)

QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAla 37
DB 133 CGCATCCGAGCAAGTTTCGAGCAGTTTCATCAGACAGATGCAGACCTCAGAGGGGCTC 192

QY 38 ---ValAlaLeuThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAsp 56
DB 193 GACCTGAACTGAAAGCAGCGAGCGGTGAGACATGGCATGTGGTGTGAGCAAGGTTGCC 252

QY 57 AsnLysValTrpPheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIle 76
DB 253 AATGAGCTGTTTGAAGGTGAGATGGGAGGTTTCGTAAGGCTCATGAAGTGCAGAG 312

QY 77 GlyTyrLeuLeuIlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsn 96
DB 313 AATGACCTCTTGCTTCATCATGCATGGCAATGCTCTTCGAGGTCTTAATCTTCGAC 372

QY 97 LeuSerHisSerGlu---IleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsn 115
DB 373 CCGAGTGGCTGCAGAAATGCTCTCTCTTCGCGGAGCAAGTGGCAACACTTTCAC 432

QY 116 HisPheLysArgAlaArgLeuPheGluAspLeuGluAspAlaGluValIlePhe 135
DB 433 AATATGTTGGTCAAGCGCGGTGAGCAATACCTGATTCGAT---GATACTAGGGTG 489

QY 136 ProSerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysGlyTyr 155
DB 490 CCGTCTCAGTTTGTGTGATCCCTCAAAAGGCTCTTACTTCAAGAAATGCGAGTGGC--- 546

QY 156 AlaSerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluLeuProThrPro 175
DB 547 -----AAAATAAATCAAGCAAGAG 567

QY 176 ThrProLysIleProLysLysArgGlyArgLysLysAsnAlaAspProGluGluIle 195
DB 568 CTTCCTGAATCACCAGCAGCAGCAGCTGTCATGTTAAAGCTTGAGCAACTGAGCAACAG 627

QY 196 AsnSerSerAlaProArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAla 215
DB 628 AAGAGT-----GATGATGACACATGACTGACCGCGTTACTTACTTTCGAGG 675

QY 216 SerAlaArgLysArgThrValThrAlaGluGluArgGluAlaIleAsnAlaLys 235
DB 676 ACTGCAAGTCAGTAAAT-----GAGGATGAAAAACAGGAATATCATGGAGCGTGT 726

QY 236 ThrPheGluProThrAsnProPhePheArgValValLeuArgProSerTyrLeuTyrArg 255
DB 727 TCCATCCGACCGGGGAATCCGCGAATCGTGTGTGTTCTGCTGATGACCCCATCTCCACGC 786

QY 256 ---GlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeu---SerGlyIle 273
DB 787 AGAACAACTTCTGACCACTCCCAAGTAAATTTGCGCTGATCATCTTCAGAGCAAGCG 846

QY 274 SerGlyPheIleLysValGlnLeuAla---GlnLysGlnTrpProValArgCysLeuTyr 292
DB 847 CGTGAAGTCTCTGCTCTAGCTGAGCAGAGAACAGCAAGTGGCATGTCAGTCTCTACTAT 906

QY 293 LysAlaLysArgAlaLysPheSer---GlnGlyTyrTrpTyrGluPheThrLeuGluAsn 311
DB 907 TCCAGCAGGACCATGCTCTCAACTGCCAAGCTTTGGGTCAAGTTTGTGAGGACCAATGGG 966

RESULT 14
ADC03415
ID ADC03415 standard; DNA, 1041 BP.
XX
AC ADC03415;
XX
DT 18-DEC-2003 (first entry)
XX
DE Rice flowering time-related gene #28.
XX
KW flowering time; rice; flower architecture; plant biotechnology;
KW commercial plant farming; agriculture; flowering-related protein; gene;
KW ds.
XX
OS Oryza sativa.
XX
PN WO2003000904-A2.
XX
PD 03-JAN-2003.
XX
PF 24-JUN-2002; 2002WO-EP006968.
XX
PR 22-JUN-2001; 2001US-0300112P.
PR 26-SEP-2001; 2001US-0325277P.
PR 30-NOV-2001; 2001US-0334984P.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PA Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T;
PI Glazebrook J, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX
XX WPI; 2003-229340/22.
DR P-PSDB; ADC03416.
XX
XX New isolated nucleic acid molecule encoding a polypeptide modulating
PT flower architecture and flowering time, useful in the area of plant
PT biotechnology, and commercial plant farming and agriculture.
XX
PS Claim 4; SEQ ID NO 55; 323pp; English.
XX
CC The invention comprises the amino acid and coding sequences of proteins
CC involved in the control of flowering time in rice. The DNA and protein
CC sequences of the invention are useful for modulating flower architecture
CC and flowering time, the DNA and protein sequences are useful in the area
CC of plant biotechnology, commercial plant farming and agriculture. The
CC present DNA sequence encodes a rice flowering-related protein of the
CC invention.
XX
SQ Sequence 1041 BP; 303 A; 218 C; 261 G; 259 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,83e-20 Length: 1041
Score: 293.50 Matches: 88
Percent Similarity: 45.43% Conservative: 56
Best Local Similarity: 27.76% Mismatches: 148
Query Match: 16.34% Indels: 25
DB: 10 Gaps: 11

US-10-088-187A-11 (1-341) x ADC03415 (1-1041)

QY 20 LeuArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAla 39
DB 58 ATTAGCATACCGGAGAAAGTTGCGGACAGATTGAGTTCAGATCACCAGAGGTTCAAC 117

QY 40 LeuThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaPheAsnLysIle 59
DB 118 CTGAAGCACCACCGATGCGGAAACATGCGGTGTCAGCGTCCGAAAGGTTGCTGATGAGCTG 177

QY 60 TrpPheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeu 79

```
Db 178 ATCTCTCATGTCAGGATGGAGGATTTTGGCAAGGCTCATGAATTCGAGGAGATGACCTC 237
Qy 80 LeuilePhe-----ArgTyrGluGlyAenSerAlaPheSerValTyrilePhe 95
Db 238 CTGTCTTCTCAGTTCGCAATGGCGTTCATGAGGAGCTCTCTTTGATGATCTAATCTTT 297
Qy 96 AenLeuSerHisSerGlu---TleAenTyrHisSerThrGlyLeuMetAenSerAlaHis 114
Db 298 GATGCAAGTGGCTGTGAGAAAGTCTCTGCTCTTTATTGGTAAAGAACACAGTATAGT 357
Qy 115 AenHisPheGlyArgAlaGluPheGluAenLeuGluAenGluAenGluValle 134
Db 358 GGTGAGACTCAGAAATGCTGAGCAATACCAATTAAGTATCTCAAGATACAGCA--- 414
Qy 135 PheProSerSer---ValTyrProSerProLeuProGluSerThrValProAlaenLys 153
Db 415 ---CCATCAATCTTCTGATGCTCTCAAGGCTCTCTCTCAAGAAATGAAAT 471
Qy 154 GlyTyrAlaSerSerAlaileGlnThrLeuPheThrGlyProValLysAlaGluPro 173
Db 472 GCGAAGACCAAG-----ACTAATCCAAATAAA---GATGATCCA 507
Qy 174 ThrProThrProLysIleProLysLysArgGlyArgLysLysAenAlaPheProGlu 193
Db 508 TTTTCAACAGGAAGAGGCTGAAGATCCAAACAGCAGCGCTCTCATGTCAAACATGAG 567
Qy 194 GluileAenSerSerAlaProArgAspAspProGlu-----AsnArgSerLys 210
Db 568 ATGATTCAGGAAGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 627
Qy 211 PheTyrGluSerAlaSerAlaArgLysArgThrValThrAlaGluGluArgGluGala 230
Db 628 TACTACTACTCGAGGTTTGGC-----AATTACCTAAGTGGGAGGAGTGAAGAG--- 678
Qy 231 IleAenAlaLysThrPheGluProThrAenProPheArgValValLeuArgPro 250
Db 679 ATATTCAAGTTGGTGTGCTCACTGCAACAGGCAATCTGTATTGTGACTGTTCTGCGAGCG 738
Qy 251 SerTyrLeuTyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeu 270
Db 739 CCACAAATTCATCGCAAGGCTCTGATGCTCCCGAGGATTTGACGCTGATCATCTT 798
Qy 271 SerGlyIleSerGlyPheIleLysValGlnLeuAlaGluLys-----GlnTrpProVal 288
Db 799 GACGAGATCACAGAAATCTGCTGTGATGAGACCAACAAAGAGGAGAAATGGTATGTC 858
Qy 289 ArgCysLeuTyrLysAlaGlyArgAlaLysPheSer---GlnGlyTyrTrpGluPheThr 307
Db 859 AAATACTACATGCAAGCACTACCGAGGCTTCAATTCGCAACGCTGATCAAGTTTCATC 918
Qy 308 LeuGluAenAenLeuGlyGluGlyAspValCysValPheGluLeuLeuArg 324
Db 919 CCGGAGAACAGGCTCGGAGAGGGTACATCTGATCTTCGAGCTGATGAAA 969

RESULT 15
AAC35159
ID AAC35159 standard; DNA; 796 BP.
XX
AC AAC35159;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9203.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
FN BP1033405-A2.
XX
```

PD 06-SEP-2000.

XX 25-FEB-2000; 200BP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 23-MAR-1999; 99US-0123548P.

XX 25-MAR-1999; 99US-0125788P.

XX 29-MAR-1999; 99US-0126264P.

XX 01-APR-1999; 99US-0126785P.

XX 06-APR-1999; 99US-0127462P.

XX 08-APR-1999; 99US-0128234P.

XX 16-APR-1999; 99US-0128714P.

XX 19-APR-1999; 99US-0129845P.

XX 21-APR-1999; 99US-0130077P.

XX 23-APR-1999; 99US-0130449P.

XX 28-APR-1999; 99US-0130510P.

XX 30-APR-1999; 99US-0130891P.

XX 30-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 04-MAY-1999; 99US-0132407P.

XX 05-MAY-1999; 99US-0132484P.

XX 06-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 07-MAY-1999; 99US-0132487P.

XX 11-MAY-1999; 99US-0132863P.

XX 14-MAY-1999; 99US-0132866P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 14-MAY-1999; 99US-0134370P.

XX 18-MAY-1999; 99US-0134768P.

XX 19-MAY-1999; 99US-0134941P.

XX 20-MAY-1999; 99US-0135124P.

XX 21-MAY-1999; 99US-0135353P.

XX 24-MAY-1999; 99US-0135629P.

XX 25-MAY-1999; 99US-0136021P.

XX 27-MAY-1999; 99US-0136392P.

XX 28-MAY-1999; 99US-0136782P.

XX 01-JUN-1999; 99US-0137222P.

XX 03-JUN-1999; 99US-0137528P.

XX 04-JUN-1999; 99US-0137502P.

XX 07-JUN-1999; 99US-0137724P.

XX 08-JUN-1999; 99US-0138094P.

XX 10-JUN-1999; 99US-0138540P.

XX 10-JUN-1999; 99US-0138847P.

XX 14-JUN-1999; 99US-0139119P.

XX 16-JUN-1999; 99US-0139452P.

XX 16-JUN-1999; 99US-0139453P.

XX 17-JUN-1999; 99US-0139454P.

XX 18-JUN-1999; 99US-0139455P.

XX 18-JUN-1999; 99US-0139456P.

XX 18-JUN-1999; 99US-0139457P.

XX 18-JUN-1999; 99US-0139458P.

XX 18-JUN-1999; 99US-0139459P.

XX 18-JUN-1999; 99US-0139460P.

XX 18-JUN-1999; 99US-0139461P.

XX 18-JUN-1999; 99US-0139462P.

XX 18-JUN-1999; 99US-0139463P.

XX 18-JUN-1999; 99US-0139750P.

XX 21-JUN-1999; 99US-0139763P.

XX 22-JUN-1999; 99US-0139817P.

XX 23-JUN-1999; 99US-0139899P.

XX 23-JUN-1999; 99US-0140353P.

XX 23-JUN-1999; 99US-0140354P.

XX 24-JUN-1999; 99US-0140695P.

XX 28-JUN-1999; 99US-0140823P.

XX 29-JUN-1999; 99US-0140991P.

XX 30-JUN-1999; 99US-0141287P.

XX 01-JUL-1999; 99US-0141842P.

XX 01-JUL-1999; 99US-0142154P.

XX 02-JUL-1999; 99US-0142055P.

Page 22

Page 22

```
QY 126 LeuGluAspGluAspAlaGluValIlePheProSerSerValTyrProSerProLeuPro 145
Db ::: ||| ||| ::::|||
428 ATCATCGACACGCGCAGATGATATAATC----- 454
QY 146 GluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThr 165
Db ----- 454
QY 166 GlyProValLysAlaGluProThrProThrProLysIleProLysLysArgGlyArg 185
Db ||| ||| ||| ::|
455 -----GAAATTACTGGCAAGAGTTTGTGGAACTCAGGGTACAAGG 496
QY 186 LysLysLys-AsnAlaAspProGluGluIleAsnSerSerAlaPro---ArgAspAspAs 204
Db ||| ||| ||| ||| ::|
497 AAAAATAAATCAATCTCTTAAATGGTGTGTATACAGACACAAATCCAAAGAGAGACCAAGA 556
QY 204 pProGluAsnArgSerLysPheTyr-----GluSerAlaSerAlaAr 218
Db : ::| ::| ::| ::| ::| ::|
557 GACATCGAGTTTGACAAAGATCTACACGATGTTGATGTCAGTGTCTTAAAGAGAA 616
QY 218 gLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThr 236
Db :|| ::| ::| ::| ::| ::|
617 GAAGAAGACAAGAGAGTCTCTCAGGGGATACAGAGCTTTTAACAGACCGAGACC 671
```

Search completed: December 30, 2004, 06:32:39
Job time : 581 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 30, 2004, 06:16:43 ; Search time 218 Seconds
(without alignments)
1111.830 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796

Sequence: 1 MRPFPFKLIFSSTIQKRL.....LLRTRDFVLKVTAFRVNEYV 341

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model -DEV=xlh
-Q/cgn2_1/USFTO_spool/US1008187/runat_29122004_115843_147/app_query.fasta_1.519
-DB=issued Patents NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1008187@cgn1.1.69@runat_29122004_115843_147 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	5.8	2269	3	US-09-351-457-3
2	104	5.8	2269	3	US-09-561-500-3
3	104	5.8	2269	3	US-09-561-108-3
4	104	5.8	2269	4	US-09-351-543-3
5	104	5.8	2269	4	US-09-561-526-3
6	104	5.8	2269	4	US-09-561-499-3
7	104	5.8	2269	4	US-09-998-831-3
8	104	5.8	2269	4	US-09-561-005-3
9	104	5.8	2269	4	US-08-819-386-3
10	104	5.8	2282	1	US-08-373-579-5
11	104	5.8	2282	1	US-08-418-595-5
12	104	5.8	2282	2	US-08-665-926-5

13	104	5.8	2282	3	US-09-162-437-5	Sequence 5, Appl1
14	104	5.8	2282	3	US-08-740-223A-5	Sequence 5, Appl1
15	104	5.8	2282	4	US-08-817-318-5	Sequence 5, Appl1
16	104	5.8	2282	4	US-09-709-188-5	Sequence 5, Appl1
17	104	5.8	2282	4	US-09-442-717-5	Sequence 5, Appl1
18	104	5.8	2282	4	US-09-689-020-5	Sequence 5, Appl1
19	97.5	5.4	3537	4	US-09-543-681A-3647	Sequence 3647, Ap
20	97.5	5.4	5751	4	US-09-023-655A-1415	Sequence 1415, Ap
21	97.5	5.4	11907	3	US-08-061-376-4	Sequence 4, Appl1
22	97.5	5.4	14255	1	US-08-320-559-1	Sequence 1, Appl1
23	97.5	5.4	14255	1	US-08-327-392-1	Sequence 1, Appl1
24	97.5	5.4	14255	1	US-08-306-691B-55	Sequence 55, Appl1
25	97.5	5.4	14255	3	US-08-545-860D-1	Sequence 1, Appl1
26	97.5	5.4	14255	5	PCT-US94-04496-1	Sequence 1, Appl1
27	97	5.4	1200	4	US-09-107-532A-2744	Sequence 2744, Ap
28	96.5	5.4	4899	4	US-09-799-451-714	Sequence 714, Appl
29	95.5	5.3	3226	3	US-08-539-205A-5	Sequence 5, Appl1
30	95.5	5.3	3226	4	US-09-392-163A-5	Sequence 5, Appl1
31	94.5	5.3	1494	4	US-09-248-796A-1806	Sequence 1806, Ap
32	94.5	5.3	2082	4	US-09-248-796A-12787	Sequence 12787, A
33	94.5	5.3	6156	4	US-10-204-708-60	Sequence 60, Appl
34	93.5	5.2	825	4	US-09-489-039A-3190	Sequence 3190, Ap
35	92	5.1	2247	4	US-09-270-767-12793	Sequence 12793, A
36	92	5.1	71989	3	US-09-443-501A-2	Sequence 2, Appl1
37	91.5	5.1	3661	4	US-09-762-481B-1	Sequence 1, Appl1
38	91.5	5.1	7888	4	US-09-919-497-48	Sequence 48, Appl
39	91.5	5.1	8312	4	US-09-620-312D-1048	Sequence 1048, Ap
40	91	5.1	1420	1	US-08-080-255-3	Sequence 3, Appl1
41	91	5.1	1420	3	US-08-463-713-3	Sequence 3, Appl1
42	91	5.1	1420	5	PCT-US93-05857-3	Sequence 3, Appl1
43	91	5.1	3131	3	US-09-035-648-23	Sequence 23, Appl
44	91	5.1	3131	3	US-09-001-951-23	Sequence 23, Appl
45	91	5.1	3131	4	US-08-818-829-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1

US-09-351-457-3

; Sequence 3, Application US/09351457

; Patent No. 6312694

; GENERAL INFORMATION:

; APPLICANT: THORPE, PHILIP E.

; APPLICANT: RAN, SOPHIA

; TITLE OF INVENTION: CANCER TREATMENT METHODS USING THERAPEUTIC CONJUGATES

; TITLE OF INVENTION: THAT BIND TO AMINOPHOSPHOLIPIDS

; FILE REFERENCE: 4001.002300

; CURRENT APPLICATION NUMBER: US/09/351,457

; CURRENT FILING DATE: 1999-07-12

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 2269

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-351-457-3

Alignment Scores:

Pred. No.: 0.0263 Length: 2269

Score: 104.00 Matches: 58

Percent Similarity: 37.55% Conservative: 34

Best Local Similarity: 23.67% Mismatches: 92

Query Match: 5.79% Indels: 61

DB: 3 Gaps: 10

US-10-088-187A-11 (1-341) x US-09-351-457-3 (1-2269)

QY 45 GlyHisValTrpArgValGlyLeuArgGlyValAlaAspAsnLysIleTrpPheGlnAspGly 64

Db 1304 GGAGCGGGTGGACAAATTTTCAGCGAGTGGCGGCTTTTCAGAGGACT 1363

QY 65 TrpGln-----Glu 67

Db	1364	TGGAAGAAATATAAAGTGGGATTTGGTTAAACCCCTTCAGGAGAAATATTGGCTGGAAATGAG	1423
Qy	68	PheValAspArgTyrSerIleAArgIleGlyTyrLeuLeu-----IlePheAArgTyr	84
Db	1424	TTTTGTTTCGCAACTGACTAATACGACACGCTATGTGCTTAAATATACACCTTAAAGACTGG	1483
Qy	85	GlucGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn	103
Db	1484	GAAGGGAATGAGGCTTACTCATTTGTATGAACATTTCTTATCTCTCAAGTGAAGAACTCAAT	1543
Qy	104	Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysAArgAlaArg	121
Db	1544	TATAGGATTCACCTTAAAGGACTTACAGGAGACCGCGCAAAATAAGCACCATCAGCCAA	1603
Qy	122	LeuPheGluAspLeuGluAspGluAspAlaGlu-----ValIlePheProSer	137
Db	1604	CCAGGAAATGATTTTAGCACAAAGGATGGAGACAAACGACAAATGATTGTCAAATGTTC	1663
Qy	138	SerValTyr-----ProSerProLeu-ProGluSe	147
Db	1664	CAAAATGCTAAACAGGAGGCTGGTGGTTGTGATGCATGTGGTCTCTTCCAACTTTGAACGGGAATG	1723
Qy	147	rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr	167
Db	1724	TACTATCCACAGGCGAGACACAAATAAGTTCACCGGCATTAATGGTACTACTCGAAA	1783
Qy	167	ovalLysAlaGluGluProThrPro-----LysIleProLysLy	182
Db	1784	GGCTCAGGCTATTTCGCTCAAGGCCACCAACCATGATGATCCGACGACAGAT	1835
Qy	182	sArgGlyArgIlyLysLysAsnAlaAppProGluIleAsnSerSerAlaProArgAs	202
Db	1836	-----TCYAAACATCCAGTCCACCTGAGAACTGCTCGAACTATTTCAAAGA	1885
Qy	202	pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa	222
Db	1886	CTTAAGCCCACTGTCAC-----TGAAGTCAACGGCTGGCGCACTGT	1924
Qy	222	lThrAlaGluGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh	237
Db	1925	GTCTCTTCCACACAGAGGGCGTGTGCTCGTGCTCAGGGACCCACATGTCTCCAGATT	1984
Qy	237	eGluProThrAsn	241
Db	1985	AGACCTGTAAAC	1997

RESULT 2

US-09-561-500-3
; Sequence 3, Application US/09561500
; Patent No. 6342219
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe

: APPLICANT: Rolf A. Brekken
 : TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
 : FILE REFERENCE: 4001.002500
 : CURRENT APPLICATION NUMBER: US/09/561,500
 : CURRENT FILING DATE: 2000-04-28
 : PRIOR APPLICATION NUMBER: 60/131,432
 : PRIOR FILING DATE: 1999-04-28

RESULT 3

```

US-09-561-108-3
; Sequence 3, Application US/09561108
; Patent No. 6342221
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE C
; FILE REFERENCE: 4001-002584
; CURRENT APPLICATION NUMBER: US/09561,108
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2269

```

; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-561-108-3

Alignment Scores:
 Pred. No.: 0.0263 Length: 2269
 Score: 104.00 Matches: 58
 Percent Similarity: 37.55% Conservatve: 34
 Best Local Similarity: 23.67% Mismatches: 92
 Query Match: 5.79% Indels: 61
 DB: 10 Gaps: 10

US-10-088-187A-11 (1-341) x US-09-561-108-3 (1-2269)

```

QY 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
Db 1304 GGAGCGCGGTGGACAAATTATTTCAGCGACGTGAGGATGGCAGCGTGTGATTTCAGAGGACT 1363
QY 65 TrpGln-----Glu 67
Db 1364 TGGAAAGAAATATAAAGTGGGATTTGGTAAACCCCTTCAGGAGAAATATTGGCTGGGAAATGAG 1423
QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
Db 1424 TTTGTTTCGCAACTGACTAATCAGCAACGCTATGTCTTAAATATACACCTTAAAGACTGG 1483
QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db 1484 GAAGGGAATGAGCTTACTCTATGTATGAACATTTCTCTCTCAAGTGAAGAACTCAAT 1543
QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
Db 1544 TATAGGATTCACCTTAAAGGACTTACAGGACAGCGGCAAAATAAGCAGCATCAGCCAA 1603
QY 122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
Db 1604 CCAGGAAATGATTTTAGCAAAAGGATGGAGACACGACAAATGATTTTGCAAAATGTTCA 1663
QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147
Db 1664 CAATGCTAACAGGAGCGTGGTGGTTTGATGATGTCATGTGGTCTTCCAACTTGAACGGAATG 1723
QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db 1724 TACTATCCACAGGCGAGACACAAATAGTTCAACGGCATTAATGGTACTACTGGAAA 1783
QY 167 oValLysAlaGluGluProThrProThrPro-----LysIleProLysLy 182
Db 1784 GGCTCAGGCTATTTCGCTCAAGGCCACACCATGATGATCCGACGACAGATT----- 1835
QY 182 sArgGlyArgLysLysLysAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db 1836 -----TCTAAACATCCAGTCCACCTGAGGAACTGTCTCGAACTATTTCGAAGA 1885
QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
Db 1886 CTTAAGCCCGAGTCAC-----TGAAGTCACGGCTCGGCACTGT 1924
QY 222 lThrAlaGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
Db 1925 GTCTCTTCCACACAGAGGGGTGTGCTCGTGTGCTGACGGGACCCACATGCTCCAGATT 1984
QY 237 eGluProThrAsn 241
Db 1985 AGAGCCTGTAAAC 1997
  
```

RESULT 4

US-09-351-543-3
 ; Sequence 3: Application US/09351543
 ; Patent No. 6406693
 ; GENERAL INFORMATION:
 ; APPLICANT: THORPE, PHILIP E.
 ; APPLICANT: RAN, SOPHIA

; TITLE OF INVENTION: CANCER TREATMENT METHODS USING ANTIBODIES TO
 ; FILE OF INVENTION: AMINOPHOSPHOLIPIDS

; FILE REFERENCE: 4001.002200
 ; CURRENT APPLICATION NUMBER: US/09/351,543
 ; CURRENT FILING DATE: 1999-07-12
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 2269
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-351-543-3

Alignment Scores:

Pred. No.: 0.0263 Length: 2269
 Score: 104.00 Matches: 58
 Percent Similarity: 37.55% Conservatve: 34
 Best Local Similarity: 23.67% Mismatches: 92
 Query Match: 5.79% Indels: 61
 DB: 10 Gaps: 10

US-10-088-187A-11 (1-341) x US-09-351-543-3 (1-2269)

```

QY 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
Db 1304 GGAGCGCGGTGGACAAATTATTTCAGCGACGTGAGGATGGCAGCGTGTGATTTCAGAGGACT 1363
QY 65 TrpGln-----Glu 67
Db 1364 TGGAAAGAAATATAAAGTGGGATTTGGTAAACCCCTTCAGGAGAAATATTGGCTGGGAAATGAG 1423
QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
Db 1424 TTTGTTTCGCAACTGACTAATCAGCAACGCTATGTCTTAAATATACACCTTAAAGACTGG 1483
QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db 1484 GAAGGGAATGAGCTTACTCTATGTATGAACATTTCTCTCTCAAGTGAAGAACTCAAT 1543
QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
Db 1544 TATAGGATTCACCTTAAAGGACTTACAGGACAGCGGCAAAATAAGCAGCATCAGCCAA 1603
QY 122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
Db 1604 CCAGGAAATGATTTTAGCAAAAGGATGGAGACACGACAAATGATTTTGCAAAATGTTCA 1663
QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147
Db 1664 CAATGCTAACAGGAGCGTGGTGGTTTGATGATGTCATGTGGTCTTCCAACTTGAACGGAATG 1723
QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db 1724 TACTATCCACAGGCGAGACACAAATAGTTCAACGGCATTAATGGTACTACTGGAAA 1783
QY 167 oValLysAlaGluGluProThrProThrPro-----LysIleProLysLy 182
Db 1784 GGCTCAGGCTATTTCGCTCAAGGCCACACCATGATGATCCGACGACAGATT----- 1835
QY 182 sArgGlyArgLysLysLysAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db 1836 -----TCTAAACATCCAGTCCACCTGAGGAACTGTCTCGAACTATTTCGAAGA 1885
QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
Db 1886 CTTAAGCCCGAGTCAC-----TGAAGTCACGGCTCGGCACTGT 1924
QY 222 lThrAlaGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
Db 1925 GTCTCTTCCACACAGAGGGGTGTGCTCGTGTGCTGACGGGACCCACATGCTCCAGATT 1984
QY 237 eGluProThrAsn 241
  
```

```
Db      1985 AGAGCCTGTAAC 1997
RESULT 5
US-09-561-526-3
; Sequence 3, Application US/09561526
; Patent No. 6416758
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002586
; CURRENT APPLICATION NUMBER: US/09/561,526
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-561-526-3
Alignment Scores:
Pred. No.: 0.0263      Length: 2269
Score: 104.00         Matches: 58
Percent Similarity: 37.55%      Conservative: 34
Best Local Similarity: 23.67%    Mismatches: 92
Query Match: 5.79%             Indels: 61
DB: 4                     Gaps: 10

US-10-088-187A-11 (1-341) x US-09-561-526-3 (1-2269)
Qy      45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
Db      1304 GGAGCGGGTGGACAATTATTACGCGACGTGAGGATGGCGGTGATTTTCAGAGGACT 1363
Qy      65 TrpGln-----
Db      1364 TGGAAAGAATATAAAGTGGGATTTGGTAACCTTCAGGAGAATATTGGCTGGGAATGAG 1423
Qy      68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
Db      1424 TTGTGTCGAAGTACTGACTTAATCAGCAACGCTATGCTTAAATACACCTTAAAGACTGG 1483
Qy      85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db      1484 GAAGGAATGAGGCTTACTCATTTGTATGAACATTTCTATCTCTCAAGTGAAGAACTCAAT 1543
Qy      104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
Db      1544 TATAGGATTCACCTTAAAGGACTTACAGGACGCGGCAAAATAAGCAGCATCGACCAA 1603
Qy      122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
Db      1604 CAAATGCTAACAGGAGGCTGCTGTTGATGATGCTGCTTCCAACTTGAACGGAATG 1723
Qy      147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db      1724 TACTATCCACAGGACGACACAAATAAGTTCAACGGCATTAATAGTGTACTACTGGAA 1783
Qy      167 oValLysAlaGluGluProThrProThrPro-----LysIleProLysIly 182
Db      1784 GGCTCAGGCTATTTCGCTCAAGGCCACACCATGATCGACACGAGCAT-----1835
Qy      182 sArgGlyA-rGlyLysAlaAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db      1836 -----TCTAAACATCCAGCTCCAGCTGAGGAAGTGTCTCGAACTATTTCGAAGA 1885

US-10-088-187A-11 (1-341) x US-09-561-499-3 (1-2269)
Qy      45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
Db      1304 GGAGCGGGTGGACAATTATTACGCGACGTGAGGATGGCGGTGATTTTCAGAGGACT 1363
Qy      65 TrpGln-----
Db      1364 TGGAAAGAATATAAAGTGGGATTTGGTAACCTTCAGGAGAATATTGGCTGGGAATGAG 1423
Qy      68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
Db      1424 TTGTGTCGAAGTACTGACTTAATCAGCAACGCTATGCTTAAATACACCTTAAAGACTGG 1483
Qy      85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db      1484 GAAGGAATGAGGCTTACTCATTTGTATGAACATTTCTATCTCTCAAGTGAAGAACTCAAT 1543
Qy      104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
Db      1544 TATAGGATTCACCTTAAAGGACTTACAGGACGCGGCAAAATAAGCAGCATCGACCAA 1603
Qy      122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
Db      1604 CAAATGCTAACAGGAGGCTGCTGTTGATGATGCTGCTTCCAACTTGAACGGAATG 1723
Qy      147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db      1724 TACTATCCACAGGACGACACAAATAAGTTCAACGGCATTAATAGTGTACTACTGGAA 1783
Qy      167 oValLysAlaGluGluProThrProThrPro-----LysIleProLysIly 182
Db      1784 GGCTCAGGCTATTTCGCTCAAGGCCACACCATGATCGACACGAGCAT-----1835
Qy      182 sArgGlyA-rGlyLysAlaAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db      1836 -----TCTAAACATCCAGCTCCAGCTGAGGAAGTGTCTCGAACTATTTCGAAGA 1885
```

Db	1724	TACTATCCACAGGCGACACAAATAAGTTCAACGGCATTAATAAGTTACTACTACTGAAA	1783
Qy	167	oValLyeAlaGluProThrProThrPro-----LysileProLysLy	182
Db	1784	GGCTCAGGCTATTCCGCTCAAGGCCCAACCATGATGATCCACGACAGATT-----	1835
Qy	182	sArgGlyArgLysLysLysAsnAlaApeProGluLuiLeAsnSerSerAlaProAaGAs	202
Db	1836	-----TCYAAACATCCAGTCACCTCGAGAAACTGTCTCGAACTATTTCYCAANGA	1885
Qy	202	pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa	222
Db	1886	CTTAAGCCCACTGTCAC-----TGAAGAGTCACGGCTGGCGCACTGT	1924
Qy	222	lThrAlaGluGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh	237
Db	1925	GTCTCTTCACACACAGAGCGGTGTGCTCGTGCTCAGCGGACCCACATGCTCCAGATT	1984
Qy	237	eGluProThrAsn	241
Db	1985	AGAGCCTGTAAC	1997

RESULT 7

```

US-09-998-831-3
; Sequence 3, Application US/09998831
; Patent No. 6676941
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY
; TITLE OF INVENTION: INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/998,831
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 09/561,108
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-831-3

```

Alignment Scores:		
Pred. No.:	0.0253	2269
Score:	104.00	Matches: 58
Percent Similarity:	37.5%	Conservative: 32
Best Local Similarity:	23.67%	Mismatches: 92
Query Match:	5.79%	Indels: 61
DB:	4	Gaps: 10

US-10-088-187A-11 (1-341) X US-09-998-831-3 (1-2269)

Qy	45	GlyHisValTrrArgValGlyLeuArgLysAlaAspAsnLysIleTrppheGlnAspGly	64
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Dd	1304	GGAGCGCGGTGGCAAAATATTTCAGCGACGTGCAGGATGCCAGCTTGATTTTCAGAGGACT	1363
Qy	65	TripGln-----	77
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Dd	1364	TGGNAAGAATAAAGTCGGATTGGTTAACCC TTCAGGAGANAATTGGCTGGGMAATGAG	1423
Qy	68	PheValAspArgTyrrSerIleArgIleGlyTyrrLeuLeu-----IlePheArgrTyrr	84
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Dd	1424	TTTTGTTTCGCAACTGCATAATCAGACAAGCTATGTGCTTAAAATACACCTTTAAAGACTGG	1483
Qy	85	GluCluIAsnSerAlaPheSerValTyrr--IlePheAsnLeuSerHisSerGluIleAsn	103
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Dd	1484	GAAGGGAATGAGGGTTACTCATGTATGAACATTTCATCTCTCAAGTGAAGAACTCAAT	1543
Qy	104	Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg	121
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Dd	1544	TATAGGATTCACCTTAAGAGACTTACAGGAGACACCGCGCAAAAATAAGCAGCATCAGCCAA	1603

Qy	122	LeuPheGluAspLeuGluAspGluAspAlaGlu-----ValIlePheProSer	137
		:::	
Db	1604	CCAGGAATGATTTTAGCACAAGAGGATGGAGACAACGACAAATGTATTTGCAAATGTTCA	1663
		:::	
Qy	138	SerValTyr-----ProSerProLeu-ProGluSe	147
		:::	
Db	1664	CNAATGCTAACAGGAGGCTGGTGGTTTGATGTCATGTGGTCTCTTCCAACTTGAACGGGATG	1723
		:::	
Qy	147	rThValProAlaAsnIysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr	167
		:::	
Db	1724	TACTATCCAGAGGCGACAGACACAAAATAAGTTCAACGGCATTAATAGTACTACTCTGGA	1783
		:::	
Qy	167	oValIysAlaGluProThrProThrPro-----LysIleProIySly	182
		:::	
Db	1784	GGCTCAGGCTATTTCGCTCAAGGCCAACACATGATGATCCGACGACGATTT-----	1835
		:::	
Qy	182	sArgIyArgIySlySlyAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs	202
		:::	
Db	1836	-----TCTAAACATCCAGTCCACCTGAGGAACTGTCTCGAACTATTTTCAAAGA	1885
		:::	
Qy	202	pAspAspProGluAsnArgSerIySPhetyrGluSerAlaSerAlaArgIySargThrVa	222
		:::	
Db	1886	CTTAAGCCCACTGCAC-----TGAAGATCAACGGCTGCGCACTGT	1924
		:::	
Qy	222	lThrAlaGluGluArgAlaIleAsn-----AlaAlaIySthrPh	237
		:::	
Db	1925	GTCTCTTCCACACAGAGGCGGTGTGCTCGGTGCTGACGGGACCCACATGCTCCAGATT	1984
		:::	
Qy	237	eGluProThrAsn	241
		:::	
Db	1985	AGAGCCTGTAAAC	1997

RESULT 8

```

US-09-561-005-3
; Sequence 3, Application US/09561005
; Patent No. 6703020
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE METHODS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002585
; CURRENT APPLICATION NUMBER: US/09/561,005
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-561-005-3

```

Alignment Scores:		
Pred. NO.:	0.0263	Length: 2269
Score:	104.00	Matches: 58
Percent Similarity:	37.55%	Conservative: 32
Best Local Similarity:	23.67%	Mismatches: 92
Query Match:	5.79%	Indels: 61
DB:	4	Gaps: 10

US-10-088-187A-11 (1-341) x US-09-561-005-3 (1-2269)

45	GlyHisValTrpArgValGlyLeuArgLeuAAspAsnLysIleTrpPheGlnAspGly	64
	:::	
1304	GGAGGGGGTGGCAATTATTCACGCACGGTGGAGATGATTTTCAGAGAACT	1363
65	TrpGln-----Glu	67
1364	TGGAAAGATATAAAGTGGGATTGGTTAAACCTTCAGAGAAATATTCGGCTGGGAAATGAG	1423
68	PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr	84

US-10-088-187A-11 (1-341) x US-09-819-386-3 (1-2269)

QY 45 GlyHisValTTPArgValGlyLeuArgLysAlaAspAsnLysIleThrPheGlnAspGly 64
Db 1304 GARGCGGCTGCAATATTATTCAGCGACGCTGAGGATGGCAGCTTATTTTCAGAGGACT 1363
QY 65 TrpGln-----Glu 67
Db 1364 TGAAGAATATAAGTGGGATTGTGTAACCTTCAGGAGATATTGGCTGGGAAATGAG 1423
QY 68 PheValAspArgTyrSerIleArgLleGlyTyrLeuLeu-----IlePheArgTyr 84
Db 1424 TTTGTTTCGACTGACTAATCAGCAACGCTATGTCTTAAATAACACCTTAAAGACTGG 1483
QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db 1484 GAAGGGAATGAGGCTTACTCATCTGTATGATGAACATTCTCTCAAGTGAAGAACTCAAT 1543
QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
Db 1544 TATAGGATTCACCTTAAAGGACTTACAGGACGAGCGGCAAAATAAGCAGCATCAGCCAA 1603
QY 122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
Db 1604 CCAGGAAATGATTTTAGCACAAGGATGAGACACGACAAATGATTTGCAATGTTCA 1663
QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147
Db 1664 CAATGCTAACAGGAGGCTGGTGTGTTGATGCTATGCTTCCCAACTTGAACGGAATG 1723
QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db 1724 TACTATCCACAGAGGCGAGACACAAATAAGTTCAACGGCATTTAAATGCTACTTGGAAA 1783
QY 167 oValLysAlaGluProThrPro-----LysIleProLysLy 182
Db 1784 GGCTCAGGCTATTCGCTCAAGGCCCAACCATGATGATCGACGACGAGATT----- 1835
QY 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db 1836 -----TCTAAACATCCAGTCCACCTGAGGAACTGTCTCGAACTATTTTCAAAGA 1885
QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
Db 1886 CTTAAGCCCATGTCAC-----TGAAAGTCACGCTGGCAGCTGT 1924
QY 222 lThrAlaGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
Db 1925 GTCTCTTCCACACAGAGGGCGTGTGCTGCTGACGGGACCCACATGCTCCAGATT 1984
QY 237 eGluProThrAsn 241
Db 1985 AGAGCCTGTAAC 1997

RESULT 10
US-08-373-579-5
; Sequence 5, Application US/08373579
; Patent No. 5650490
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: TIE-2 LIGAND, METHOD OF MAKING AND USES
; NUMBER OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM: Floppy disk
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible

US-09-819-386-3
; Sequence 3, Application US/09819386
; Patent No. 6783760
; GENERAL INFORMATION:
; APPLICANT: RAN, SOPHIA
; TITLE OF INVENTION: CANCER TREATMENT METHODS USING THERAPEUTIC CONJUGATES
; FILE REFERENCE: 4001.002300
; CURRENT APPLICATION NUMBER: US/09/819,386
; PRIOR FILING DATE: 2001-03-28
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-819-386-3

Alignment Scores:
Pred. No.: 0.0263 Length: 2269
Score: 104.00 Matches: 58
Percent Similarity: 37.55% Conservative: 34
Best Local Similarity: 23.67% Mismatches: 92
Query Match: 5.79% Indels: 61
DB: 4 Gaps: 10

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/373,579
; FILING DATE: 17-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,503
; FILING DATE: 09-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,492
; FILING DATE: 02-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/330,261
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/319,932
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cobert, Robert J.
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 330-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 345-7400
; TELEFAX: (914) 345-7721
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 357..1847
;
US-08-373-579-5
Alignment Scores:
Pred. No.: 0.0266 Length: 2282
Score: 104.00 Matches: 58
Percent Similarity: 37.55% Conservative: 34
Best Local Similarity: 23.67% Mismatches: 92
Query Match: 5.79% Indels: 61
DB: 1 Gaps: 10

US-10-088-187A-11 (1-341) x US-08-373-579-5 (1-2282)
QY 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
DB 1311 GGAGGCGGTGGACAATTATTTCAGCGACGTGAGGATGGCGGTGTTGATTTTCAGAGGACT 1370
QY 65 TrpGln-----Glu 67
DB 1371 TGGAAAGAAATATAAGTGGGATTGGTAACCTTCAGGAGAAATATTGCTGGGAAATGAG 1430
QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
DB 1431 TTTGTTTCGAAGTGAATTAATCAGCAACGCTATGCTTAAATACACCTTAAAGACTGG 1490
QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
DB 1491 GAAGGGAATGAGGCTTACTCATTTGTATGAACATTTCTATCTCTCAAGTGAAGAACTCAAT 1550
QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
DB 1551 TATAGATTACCTTAAGGACTTACAGGACAGCGCGCAAAATAAGCAGCATCAGCCAA 1610
QY 122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
DB 1611 CCAGGAATGATTTAGCACAAAGGATGGAGACACGACAAATGTTGCAAAATGTCA 1670
QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147
DB 1671 CAAATGCTAACAGGAGGCTGGTGGTTTGATGATGCTGCTTCAACTTGAACCGAATG 1730
QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
DB 1731 TACTATCCACAGAGCGCAGACACAATAAGTTCAAGCGCATTAATGTTACTACTCTGAAA 1790
QY 167 oValLysAlaGluGluProThrPro-----LysIleProLysLys 182
DB 1791 GGCTCAGGCTATTTCGCTCAAGGCCCAACCATGATGATCGACGACAGATT----- 1842
QY 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
DB 1843 -----TCTAAACATCCAGTCCACCTGAGGAACTGCTCGAAGTATTTTCAAGA 1892
QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
DB 1893 CTTAAGCCAGTGCAC-----TGAAAGTCACGGCTCGGCACTGT 1931
QY 222 lThrAlaGluGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
DB 1932 GTCTCTCTTCCACACAGAGGCGTGTGCTCGTGTGCTGACGGAGCCACATGCTCCAGATT 1991
QY 237 eGluProThrAsn 241
DB 1992 AGAGCCTGTAAC 2004

RESULT 11
US-08-418-595-5
; Sequence 5, Application US/08418595
; Patent No. 5814464
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: TIE-2 LIGAND, METHOD OF MAKING AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,595
; FILING DATE: 06-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,579
; FILING DATE: 17-JAN-1995
; APPLICATION NUMBER: US 08/353,503
; FILING DATE: 09-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,492
; FILING DATE: 02-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/330,261
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/319,932
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cobert, Robert J.
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 330-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 345-7400
; TELEFAX: (914) 345-7721
; INFORMATION FOR SEQ ID NO: 5:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 357..1847
US-08-418-595-5

Alignment Scores:
Pred. No.: 0.0266 Length: 2282
Score: 104.00 Matches: 58
Percent Similarity: 37.55% Conservative: 34
Best Local Similarity: 23.67% Mismatches: 92
Query Match: 5.79% Indels: 61
DB: Gaps: 10

US-10-088-187A-11 (1-341) x US-08-418-595-5 (1-2282)
QY 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrrPheGlnAspGly 64
DB 1311 GGAGCGCGGTGGACATATTATTCAGCGACGTGAGGATGGCAGCGTGTGATTTTCAGAGGACT 1370
QY 65 TrpGln-----Glu 67
DB 1371 TGGAAAGATATATAAGTGGGATTTGGTAACCTTCAGGAGAAATATTGGCTGGGAAATGAG 1430
QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
DB 1431 TTGTTTCCCACTGACTAATCAGCAGCGTATGCTTAAATACACCTTAAAGACTGG 1490
QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
DB 1491 GAAGGGAATGAGCTTACTCATTTGTATGAACATTTCTCTCAAGTGAAGAACTCAAT 1550
QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
DB 1551 TATAGGATTCACCTTAAAGGACTTACAGGAGACAGCCGGCAAAATAAGCAGCATCGCCAA 1610
QY 122 LeuPheGluAspLeuGluAspGluAspAlaGlu-----ValIlePheProSer 137
DB 1611 CCAGGAATGATTTTAGCACAAAGGATGGAGACACGACAAATGTTTGCATATGTTCA 1670
QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147
DB 1671 CAAATGCTAACAGGAGGCTGGTGGTTTGATGATGTCCTTCCCACTTGAACGGAATG 1730
QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
DB 1731 TACTATCCACAGAGGAGAACACAAATAGTTCAACGGCATTAATGGTACTACTGGAAA 1790
QY 167 oVallyAlaGluProThrProThrProThrPro-----LysIleProLysIy 182
DB 1791 GGCTCAGGCTATTTCGCTCAAGGCCACCAACCATGATGATCGACACGACAGATT----- 1842
QY 182 sArgGlyArgLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
DB 1843 -----TCTAACAATCCAGTCCACCTGAGGAACTGTCTCGAACTATTTCCTAAGA 1892
QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
DB 1893 CTTAAGCCCATGTCAC-----TGAAAGTCACGGCTGGCACTGT 1931
QY 222 lThrAlaGluGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
DB 1932 GTCCCTCTCCACACAGAGGGCGTGTGCTCGGTGCTGAGGGACCCACATGCTCCAGATT 1991
QY 237 eGluProThrAsn 241
DB 1992 AGAGCCTGTAAAC 2004
```

```

RESULT 12
US-08-665-926-5
; Sequence 5, Application US/08665926
; Patent No. 5851/97
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: TIE LIGAND-3, METHODS OF MAKING AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591-8707
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,926
; FILING DATE: 19-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Robert J. Cobert
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 330-H
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 345-7400
; TELEFAX: (914) 345-2113
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 357..1847
US-08-665-926-5

Alignment Scores:
Pred. No.: 0.0266 Length: 2282
Score: 104.00 Matches: 58
Percent Similarity: 37.55% Conservative: 34
Best Local Similarity: 23.67% Mismatches: 92
Query Match: 5.79% Indels: 61
DB: Gaps: 10

US-10-088-187A-11 (1-341) x US-08-665-926-5 (1-2282)
QY 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrrPheGlnAspGly 64
DB 1311 GGAGCGCGGTGGACATATTATTCAGCGACGTGAGGATGGCAGCGTGTGATTTTCAGAGGACT 1370
QY 65 TrpGln-----Glu 67
DB 1371 TGGAAAGATATATAAGTGGGATTTGGTAACCTTCAGGAGAAATATTGGCTGGGAAATGAG 1430
QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
DB 1431 TTGTTTCCCACTGACTAATCAGCAGCGTATGCTTAAATACACCTTAAAGACTGG 1490
QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
DB 1491 GAAGGGAATGAGCTTACTCATTTGTATGAACATTTCTCTCAAGTGAAGAACTCAAT 1550
QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
DB 1551 TATAGGATTCACCTTAAAGGACTTACAGGAGACAGCCGGCAAAATAAGCAGCATCGCCAA 1610
```


QY 122 LeuPheGluAspLeuGluAspGluAspAlaGlu-----ValIlePheProSer 137
Db 1611 CCAGGAATGATTTTAGCACAAGGATGGAGACACAAATGATTTGCAAAATGTTCA 1670
QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147
Db 1671 CAATGCTAACAGGAGCGTGGTGGTTGATGCGATGGTCCCTTCCAACTTGAACGGAATG 1730
QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db 1731 TACTATCCACAGAGGAGCAGACAAATAAGTTCACGGCATTAATGGTACTACTGAAA 1790
QY 167 oValLysAlaGluGluProThrProThrPro-----LysIleProLysLy 182
Db 1791 GGCTCAGGCTATTCGCTCAAGGCCACACCATGATGATCGACGACGATTT-----1842
QY 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db 1843 -----TCTAAACATCCCACTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 1892
QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
Db 1893 CTAAAGCCAGTGCAC-----TGAAGTCAAGCTGGCTGGCACTGT 1931
QY 222 lThrAlaGluGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
Db 1932 GTCTCTTCCACACAGAGGCGGTGTGCTCGGTGCTGCGGAGCCACATGCTCCAGATT 1991
QY 237 eGluProThrAsn 241
Db 1992 AGAGCCTGTAAC 2004
RESULT 13
US-09-162-437-5
Sequence 5, Application US/09162437
Patent No. 6166185
GENERAL INFORMATION:
APPLICANT: Davis, et al.
TITLE OF INVENTION: TIE-2 LIGAND, METHOD OF MAKING AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/162,437
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/418,595
FILING DATE: 06-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,579
FILING DATE: 17-JAN-1995
APPLICATION NUMBER: US 08/353,503
FILING DATE: 09-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,492
FILING DATE: 02-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/330,261
FILING DATE: 27-OCT-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/319,932
FILING DATE: 07-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Cobert, Robert J.
REGISTRATION NUMBER: 36,108
REFERENCE/DOCKET NUMBER: REG 330-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 345-7400
TELEFAX: (914) 345-7721
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2282 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 357..1847
US-09-162-437-5

Alignment Scores:
Pred. No.: 0.0266 Length: 2282
Score: 104.00 Matches: 58
Percent Similarity: 37.55% Conservative: 34
Best Local Similarity: 23.67% Mismatches: 92
Query Match: 5.79% Indels: 61
DB: 3 Gaps: 10

US-10-088-187A-11 (1-341) x US-09-162-437-5 (1-2282)

QY 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
Db 1311 GGAGCGCGGTGGACAATTATTTCAGCGACGTGAGGATGGCAGCGTGTGATTTTCAGAGGACT 1370
QY 65 TrpGln-----Glu 67
Db 1371 TGGAAAGAATATAAAGTGGGATTTGGTAACCCCTTCAGGAGAAATATTGGCTGGGAAATGAG 1430
QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
Db 1431 TTGTTCGCACTGCATCAATCAACGACGCTATGTGCTTAAATACACCTTAAAGACTGG 1490
QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db 1491 GAAGGGAATGAGCTTACTCATTTGATGACATTTCTCTCTCAAGTGAAGACTCAAT 1550
QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
Db 1551 TATAGGATTCACCTTAAGGACTTACAGGGACAGCGCGCAAAATAAGCAGCATCAGCCAA 1610
QY 122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
Db 1611 CCAGGAATGATTTTAGCACAAGGATGGAGACACAAATGATTTGCAAAATGTTCA 1670
QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147
Db 1671 CAATGCTAACAGGAGCGTGGTGGTTGATGCGATGGTCCCTTCCAACTTGAACGGAATG 1730
QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db 1731 TACTATCCACAGAGGAGCAGACAAATAAGTTCACGGCATTAATGGTACTACTGAAA 1790
QY 167 oValLysAlaGluGluProThrProThrPro-----LysIleProLysLy 182
Db 1791 GGCTCAGGCTATTCGCTCAAGGCCACACCATGATGATCGACGACGATTT-----1842
QY 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db 1843 -----TCTAAACATCCCACTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 1892
QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
Db 1992 AGAGCCTGTAAC 2004

Db1893CTTAAGCCAGTGCAC-----TGAAAGTCACGGCTGGCAGCTGT1931

Qy222lThrAlaGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh237

Db1932GTCTCTTCCACACAGAGGGCGGTGCTCGTGTGTGACGGGACCCACATCTCCAGATT1991

Qy237eGluProThrAsn241

Db1992AGAGCCTGTAAAC2004

RESULT 14

US-08-740-223A-5

Sequence 5, Application US/08/740223A

Patent No. 628584

GENERAL INFORMATION:

APPLICANT: Davis, et al.

TITLE OF INVENTION: Expressed Ligand - Vascular

INTERVENTION: Intercellular Signalling Molecule

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: Regeneron Pharmaceuticals, Inc.

STREET: 777 Old Saw Mill Road

CITY: Tarrytown

STATE: NY

COUNTRY: USA

ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/740,223A

FILING DATE: 25-OCT-1996

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 60/022/999

FILING DATE: 02-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Covert, Robert J

REGISTRATION NUMBER: 36,108

REFERENCE/DOCKET NUMBER: REG 333

TELECOMMUNICATION INFORMATION:

TELEPHONE: 914-345-7400

TELEFAX: 914-345-7721

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2282 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 357...1844

OTHER INFORMATION:

NAME/KEY: Human TIE-2 ligand 2

LOCATION: 1...2282

OTHER INFORMATION: from clone pbluescript KS

OTHER INFORMATION: encoding human TIE 2 ligand 2

US-08-740-223A-5

Alignment Scores:

Pred. No.: 0.0266

Length: 2282

Score: 104.00

Matches: 58

Percent Similarity: 37.55%

Conservative: 34

Best Local Similarity: 23.67%

Mismatches: 92

Query Match: 5.79%

Indels: 61

DB: 3

Gaps: 10

US-10-088-187A-11 (1-341) x US-08-740-223A-5 (1-2282)

Qy45GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly64

Alignment Scores:

Pred. No.: 0.0266 Length: 2282
Score: 104.00 Matches: 58
Percent Similarity: 37.55% Conservative: 34
Best Local Similarity: 23.67% Mismatches: 92
Query Match: 5.79% Indels: 61
DB: 4 Gaps: 10

US-10-088-187A-11 (1-341) x US-08-817-318-5 (1-2282)

```
QY 45 GlyHisValTyrArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1311 GGAGGCGGTGGACAATTATTACGCGACGTGAGGATGGCAGCGTTGATTTTCAGAGGACT 1370

QY 65 TrpGln-----Glu 67
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1371 TCGAAGAATATAAAGTGGGATTGGTAACCTTCAGGAGAATATTGGCTGGGAATGAG 1430

QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1431 TTTGTTTCGCAACTGACTAATCAGCAACGCTATGTCTTAAATAACACCTTAAAGACTGG 1490

QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1491 GAAGGGAATGAGCTTACTCATTTGATGAACATTTCTATCTCAAGTGAAGAAGCTCAAT 1550

QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1551 TATAGNTTCACCTTAAAGGACTTACAGGGCAGCGCGCAAAATAAGCAGCATCAGCCAA 1610

QY 122 LeuPheGluAspLeuGluAspGluAspAlaGlu-----ValIlePheProSer 137
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1611 CCAGGAATGATTTTAGCACAAAGGATGGAGACACACAAATGTATTGCAATGTTCA 1670

QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1671 CAATGCTAACAGGAGGTGGTGGTTTGATGATGGTCTTCCAACCTTGAACGGAATG 1730

QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1731 TACTATCCACAGGCGCAGACACAAATAAGTTCAACGGCATTAATGGTACTACTGGAAA 1790

QY 167 oValLysAlaGluGluProThrProThrPro-----LysIleProLysLy 182
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1791 GGCTCAGGCTATTGCTCAAGGCCACACCATCATGATCCGACACGAGATT----- 1842

QY 182 sArgGlyArgLysLysLysAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1843 -----TCTAAACATCCCAAGTCCACCTGAGGAAGTGTCTCGAAGCTATTTTCAAAGA 1892

QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1893 CTTAAGCCAGTGCAC-----TGAAAGTCACGGCTCGGCACTGT 1931

QY 222 lThrAlaGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1932 GTCTCTTCCACACAGAGGCGGTGCTCGGTGCTGACGGAGCCACATGCTCCAGATT 1991

QY 237 eGluProThrAsn 241
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1992 AGAGCCTGTAAC 2004
```

Search completed: December 30, 2004, 09:53:17
Job time : 227 secs

THIS PAGE BLANK (USPIS)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 30, 2004, 06:32:43 ; Search time 1535 Seconds
(without alignments)
1254.759 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796
Sequence: 1 MPRPFHKLIFSSTQKRL.....LLRTRDFVLKVTAPRVNXY 341

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4176236 seqs, 2824127955 residues

Total number of hits satisfying chosen parameters: 8352472

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q/cgn2_1/USPTO_spool/US1008187/runat_29122004.115844.192/app_query.fasta_1.519
-DB=Published Applications NA -QMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USPR=US1008187@cgn_1.1.480@runat_29122004.115844.192
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_1/ptodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
- 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1084	60.4	1471	17	US-10-767-795-1318	Sequence 1318, Ap
2	946	52.7	681	9	US-09-938-842A-273	Sequence 273, App
3	946	52.7	681	11	US-09-938-842A-273	Sequence 273, App
4	908.5	50.6	2141	16	US-10-424-599-118105	Sequence 118105, Ap
5	880	49.0	1111	17	US-10-767-795-1316	Sequence 1316, Ap
6	835.5	46.5	1760	16	US-10-424-599-97592	Sequence 97592, A
7	640	35.6	1062	17	US-10-767-795-1317	Sequence 1317, Ap
8	582.5	32.4	644	17	US-10-021-323-9015	Sequence 9015, Ap
9	512	28.5	426	9	US-09-770-423-237	Sequence 237, App
10	375.5	20.9	628	16	US-10-424-599-97591	Sequence 97591, A
11	330.5	18.4	650	17	US-10-767-795-1320	Sequence 1320, Ap
12	329	18.3	2207	18	US-10-739-930-3743	Sequence 3743, Ap
13	303	16.9	2338	18	US-10-425-115-89685	Sequence 89685, A
14	281	15.6	1477	17	US-10-437-963-89631	Sequence 89631, A
15	280	15.6	1455	16	US-10-425-114-32044	Sequence 32044, A
16	278	15.5	1209	17	US-10-437-963-54928	Sequence 54928, A
17	269	15.0	2253	17	US-10-437-963-75700	Sequence 75700, A
18	268	14.9	769	16	US-10-424-599-23571	Sequence 23571, A
19	265	14.8	602	17	US-10-021-323-7926	Sequence 7926, Ap
20	256.5	14.3	1699	18	US-10-425-115-107372	Sequence 107372, A
21	246	13.7	2214	17	US-10-437-963-101386	Sequence 101386, A
22	241.5	13.4	986	17	US-10-437-963-49771	Sequence 49771, A
23	229.5	12.8	2000	9	US-09-938-842A-2973	Sequence 2973, Ap
24	229.5	12.8	2000	11	US-09-938-842A-2973	Sequence 2973, Ap
25	214	11.9	921	17	US-10-437-963-27186	Sequence 27186, A
26	210.5	11.7	494	16	US-10-424-599-64127	Sequence 64127, A
27	206.5	11.5	1242	17	US-10-767-795-2666	Sequence 2666, Ap
28	205.5	11.4	1578	18	US-10-425-115-130216	Sequence 130216, A
29	197	11.0	4068	17	US-10-437-963-84667	Sequence 84667, A
30	195.5	10.9	4068	17	US-10-437-963-35166	Sequence 35166, A
31	193.5	10.8	593	16	US-10-424-599-36556	Sequence 36556, A
32	193	10.7	1611	17	US-10-437-963-60237	Sequence 60237, A
33	186	10.4	1611	17	US-10-437-963-93094	Sequence 93094, A
34	183	10.2	634	16	US-10-424-599-89606	Sequence 89606, A
35	178	9.9	1617	17	US-10-437-963-18071	Sequence 18071, A
36	174	9.7	1269	17	US-10-437-963-74762	Sequence 74762, A
37	171.5	9.5	1419	18	US-10-739-930-4151	Sequence 4151, Ap
38	170.5	9.5	1176	17	US-10-437-963-35940	Sequence 35940, A
39	167	9.3	1152	16	US-10-425-114-3941	Sequence 3941, Ap
40	166	9.2	946	17	US-10-437-963-54432	Sequence 54432, A
41	165.5	9.2	646	16	US-10-424-599-99156	Sequence 99156, A
42	162	9.0	996	16	US-10-424-599-73748	Sequence 73748, A
43	161.5	9.0	1059	9	US-09-938-842A-312	Sequence 312, App
44	161.5	9.0	1059	11	US-09-938-842A-312	Sequence 312, App
45	154.5	8.6	631	18	US-10-425-115-89684	Sequence 89684, A

ALIGNMENTS

RESULT 1
US-10-767-795-1318
; Sequence 1318, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 1318
; LENGTH: 1471
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C1540_1
US-10-767-795-1318

Alignment Scores:					
Pred. No.:	2,14e-127	Length:	1471		
Score:	1084.00	Matches:	222		
Percent Similarity:	72.48%	Conservative:	63		
Best Local Similarity:	60.49%	Mismatches:	44		
Query Match:	60.36%	Indels:	38		
DB:	17	Gaps:	8		
US-10-088-187A-11 (1-341) x US-10-767-795-1318 (1-1471)					
Qy	1 MetProArgProPheHisLysIstauLlePheSerSerThrIleGlnLysArgLeu	20			

D	B		202	ATGCACGCGCCCTTTTTCATAGCTATTCTCCTCCACTACTCTCCAGGACAGGAACAAC	G	281
Q	Y		21	ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu		40
D	B		262	AGGATACTCATTAACCTTTATTAGAANAATCAGGGATGAGCTTCTGTGTGCCGCTCTC		321
Q	Y		41	ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp		60
D	B		322	ACCGTTCCTGATGGTCATGTTTCGCGTGTAGGAATAAAGAAAGTTGACAAACAGGCTTG		381
Q	Y		61	PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleAcGileGlyTyLeuLeu		80
D	B		382	TTTAAGAGAGGGTTGGCAGGAATTTGCGAACGTTACTATACCAGTTGGCTACTTTTTG		441
Q	Y		81	IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer		100
D	B		442	ATTTTCAGATATGAAGGGAATTCGCGTTCCTCAGTGTAGCATATTTGATTTGTACAAC		501
Q	Y		101	GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla		120
D	B		502	GAAATAAACTATCAACAAAATAGCTCTTTGTGTGACTCAATCAAT---		558
Q	Y		121	ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr		140
D	B		559	TATCCAATTTGAAGAACTTGAAGATGAT-----GAATGTGTCTCTCCAGCG-		603
D	B					145

Db	604	-----CTGCCATAATTTTGTGGGGTCAAACTTAACGTATATAACTGGAGTGGA	654
Qy	146	GlUSeRThrValProAlaAsnLysGlyTyzAlaseRSeRaIleGlnThrLeu----	163
Db	655	GAAGTCAACCATCAGCACTTAGGGTGTAAATAACAACCTATTCGSGGTATGCCTTTT	714
Qy	164	-----PheThrGlyProValIys----AlaGluGluPro	173

Qy	174	ThrProThrProLysIleProLysValArgGlyValArgLysLysLeuAsnAlaAspProGlu	193
Dd	775	GCAGAGCTACCAAACTG--AAAACCAGGGAGAAAAAGCAGAAGTTTCGCGCTAGT	831
Qy	194	GluileAasnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrgLu	213
Dd	832	GAAGAGGATTATCTCTTGACATGAAGATGACATGCATATGGTAATAGATTTATGAA	891
Qy	214	SerAlaSerAlaAtGLysEargThrValThrAlaGluGluArgGluArgAlaIleAsnAla	233
Dd	892	AGTGCTCCGCCAGGAGAGAATTGTACTGCTGAAGAAAGACAGAGACGGATTAAATGCA	951
Qy	234	AlaLysThrPheGluProThrAsnPropPheArgValValLeuArgProSerTyrlieu	253
Dd	952	GCCAAAGCATTTGACCTACTAACCCCTTTCGCAGGGTCTCTTGCAGCGCTTATCTA	1011
Qy	254	TyrArgGlyCyfillewTyrlieuProSerGlyPheAlaGluLysTyrlieusergLyile	273
Dd	1012	TACAGGGGATGATTATGTACTTACCATCGGCTTTCTCGAGCAACAATCTAGACGGGGTT	1071
Ov	274	SerGlyPheIleLysValcInLeuAlaGlu---LysGlnTrpProValArgCysLeuTy	292

109	Db	GAACCACTGAATAAAGGTTATGGCGGTTCTACAGCCATCCAAAGCTTTTTCAAA---GAA	165
168	Qy	ValIysAlaGluGluProThrProThrsProIlysAlaArgGlyArgGlyLys	187
166	Db	TCATAAGCTTGAAGNA-----ACGCCCAAGGTACTTAAAGAGAGAGGAAGAAG	216
188	Qy	LysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAspAspProGluAsn	207
217	Db	AAGAATCCTAATCCGAGGAAGTAATCTTTCAACTCCCGGTGGAGATGACTCAGAGAAC	276
208	Qy	ArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrValThrAlaGluArg	227
277	Db	CGTCAAAAGTTCACGAGAGTGCTTCTGCTAGAAGAGAACTGTAACTGCAGAGAAAGA	336
228	Qy	GluArgAlaIleAsnAlaAlaLysThrPheGluProThrAsnProPhePheArgVal	247
337	Db	GAGAGACCGTCAATGCGAGCCAAACATTCGACCAACAAATCCTTACTTTAGAGTTGTT	396
248	Qy	LeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGlu	267
397	Db	CTCGACCATCATATCATACAGAGGTGTCATCATGTACTTGGCACTCTGGTTCGTGAG	456
268	Qy	LysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGluLysGlnTrpPro	287
457	Db	AAATACCTAAGTGGGATATCTGGTTCATCAAGCTCCAGCTCGGTGGAGAAA	516
288	Qy	ValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTyrTyrGluPheThr	307
517	Db	GTGAGTGGCTCTCAAAAGCAGGAGAGCTAAGTTTAGCCAAAGGATGGTATGAGTTCA	576
308	Qy	LeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArgThrArgAsp	327
577	Db	CTCGAGAACCAATATAGCGGAGGAGATGTATGTGTGTTGAGCTACTCAGAACTCGGAT	636
328	Qy	PheValLeuLysValThrAlaPheArgValAsnGluTyrVal	341
637	Db	TTGGTCTTCAAGTCACCGGCTTTCGTTTCGTTCAATGATGATGG	678

Qy Db

```

Db      517  GTGAGTGGCCCTCTACAAGCAGGAGAGCTAAGTTTATGCCAAGGATGGTATGAGTTTCA 576
Qy      308  LeuGluAsnAsnLeuGluGlyAspValCysValPheGluLeuLeuArgThrArgAsp 327
Db      577  CTCGAGAACAAATATAGCGGAGGAGATGTATGTGTGTTTGAGCTACTCAGAACTCGGAT 636
Qy      328  PheValLeuLysValThrAlaPheArgValAsnGluTyrVal 341
Db      637  TTCGTTCTCGAAGTCACCGCCTTTTCGTGTCATGAGTATGTG 678

RESULT 4
US-10-424-599-118105
; Sequence 118105, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 118105
; LENGTH: 2141
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_77658C.1
US-10-424-599-118105

```

```

; NUMBER
; SEQ ID
; LENGTH
; TYPE
; ORGAN
; FEAT
; OTHER
US-10-42

```

```

; CLUSTAL W (1.8.0)
; LENGTH: 2141
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_77658C.1
US-10-424-599-118105

Alignment Scores:
Pred. No.:          9,47e-105      Length:       2141
Score:             908.50         Matches:       198
Percent Similarity: 61.31%        Conservative:  46
Best Local Similarity: 49.75%     Mismatches:   95
Query Match:       50.58%         Indels:       59
DB:                16            Gaps:         8

US-10-088-187A-11 (1-341) x US-10-424-599-118105 (1-2141)

Qy      1 MetProArgPropheHisLysLeuIlePheSerGthrIleGln---GlulYysArg 19
           |||||:::|||||

```

Qy 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGln--GluLysArg 19

452 ATGCCGACCCCTCTTTCCACCAAGCTTCTTCTACCTCCACTGTCCACCAACCAACAA 511
QY
20 LeuArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValala 39
Db
512 CTGAGCTCCAGATAAATTTATGAGAAATATGGGGTGAACCTCTCCAAATGTTTACC 571
QY
40 LeuThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIle 59
Db
572 CTCTCTGTTCTGATGAGTGTCTGCGCATGTAGGATTTGAAAAAAGCAGACAAACAGTAT 631
QY
60 TrpPheGlnAspGlyTrpGlnGluPheValAspArgTrpSerIleArgIleGlyLeu 79
Db
632 TGTTCCTGATGGTGGAAAGATTTGTTCAACCTATTTCATTCGTTGGATCTTG 691
QY
80 LeuIlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHis 99
Db
692 TCGGTCCGACACCGCTCGAAAGTCCTGTCACCTGTCTGTTATTTTAAATTCGCTACT 751
QY
100 SerGluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArg 119
Db
752 TCTGAGATAAATCATCAATCACTACACGAGCAGTAATGAAGGCTTCACCTTCACAAAT 811
QY
120 Ala---ArgLeuPheGluAspLeuGluAspGluAlaGluValIle-----Phe 135
Db
812 TGTCTTAAATTTTGAAGAAATGAGGTGAAGATTCATTTGAATCTCGGATTCATCA 871
QY
136 ProSerSerValTyrProSerProLeuPro-----GluSer 147
Db
872 CCTCCACCTTCCACCAAGTTCATTCGAAATCAAGCTCTTGTGATCATGATAGATAAA 931
QY
148 ThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGly--- 166
Db
932 ATGATCCCGGAAAGAGTTTATAATACACCTCCAGCATTCGAGATCTGTTCAATGGATCT 991
QY
166 ----- 166
QY
992 AAATTAATAGCATAAACTGGGAGAGGTGTGTAATGCCCATCTTCAAGAGTGCCCAAT 1051
QY
167 -----ProValLysAlaGlu 171
Db
1052 TCACTAGATAAATCGATTGACCGAGACATPAGGACTTCAGTTTAAATGCGATCGAGTTAAA 1111
QY
172 GluProThrProThrProLysIle-----ProLysLysArgGly 184
Db
1112 AGATCTACTGAAGATTGAATTTGCTGCTTCCATCGAGAAAGGATGAAAAAACTACA 1171
QY
185 ArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAspAsp 204
Db
1172 AGAAGAAGCGAAATCAGATGGCCAGGAA-----CCCTCTGCTGGGCACGAGAGGAG 1225
QY
205 ProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrValThrAla 224
Db
1226 GTAGAATGGCTTTAGATTTTATGAAGAGTCTTCTGCAAGAAAAAAGAACTGTGACAGCT 1285
QY
225 GluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThrAsnProPhePhe 244
Db
1286 GAAGAAGAGAGAGGCTCATCAAGAGCAAAAGCATTTGAACCATTAATCCTTCTGT 1345
QY
245 ArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeuProSerGly 264
Db
1346 CGAGTTGCTCGCGGCCCTCATATTTATATAGGGGATGCATATATGATATGCGCATCATGC 1405
QY
265 PheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGlu--- 283
Db
1406 TTGTCAGAGAGCATTTGAATGAGATTTTCAAGGTTTATTAACCTCAGATCTCTAATGCT 1465
QY
284 LysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTrp 303
Db
1466 AGACAGTGGCGGTTGCTGCTGCTTTATAAGGAGGTAGAGCCAAAGTAAAGCCAGGGGTGG 1525
QY
304 TyrGluPheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeu 323
Db
1526 TTTGAATTTTCTATAGAGAAATTTTAGGGGAAGGTGATGCTGTGTTGAGCTGCTT 1585

QY 324 ArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyrVal 341
Db 1586 AGAATGAAGAAGTAGTCTGCAAGTTTACCATCTTTCATGTTACCGAGATGTG 1639
RESULT 5
US-10-767-795-1316
; Sequence 1316, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 1316
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C1540_3
US-10-767-795-1316

Alignment Scores:
Pred. No.: 1,46e-101 Length: 1111
Score: 880.00 Matches: 184
Percent Similarity: 67.81% Conservative: 33
Best Local Similarity: 57.50% Mismatches: 54
Query Match: 49.00% Indels: 49
Gaps: 7
DB:

US-10-088-187A-11 (1-341) x US-10-767-795-1316 (1-1111)

QY 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnLysArgLeu 20
Db 269 ATGCCACGCCCTTTTTCATAAGCTTATCTCTACCACTCCCAAGAAAGAACTG 328
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
Db 329 AGGATCCCAATACCTTCGTTAAGAAATTCGGGACGAACTTCTGTGTCGCGCTC 398
QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
Db 389 ACTGTTCTGACGCTCATGTTGGCGTGTAGGAATAGGAAGCCGACAAAGGTTTGG 448
QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80
Db 449 TTTCAAGGGGTTGGCAGGAGTTTCTAGAGCGGTACCATATCCGTGTGTGCTACGTACTG 508
QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
Db 509 GTTTTCAGATATGATGAATTCCTGTTTCACT----- 541
QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
Db 542 -----AAAAACC 547
QY 121 -ArgLeuPheGluAspLeuGluAspGluAspAla-----GluValIlePhe 135
Db 548 GTATCCATTTGAGAACTTGAAGATGACGAATGCAATCTCTCCGACATGCGAATTTGTT 607
QY 135 eProSerSerValTyrProSerProLeu-----ProGluSerThrValProAlaAs 152
Db 608 TTTGCGCTCTAAAGTAAATAACTGCATGAACCTGGGAGTGTGAATCAATTTTCGTGCGTT 667
QY 152 nLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGlu 172
Db 668 AAAGGGCATGAATAGCCAATCT-----TTTCAGGTGCGAGTG----- 704

US-10-767-795-1317
; Sequence 1317, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 1317
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(644)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C1540_2
US-10-767-795-1317

Alignment Scores:
Pred. No.: 5,65e-71 Length: 1062
Score: 640.00 Matches: 131
Percent Similarity: 82.49% Conservative: 15
Best Local Similarity: 74.01% Mismatches: 20
Query Match: 35.63% Indels: 11
DB: 17 Gaps: 3

US-10-088-187A-11 (1-341) x US-10-767-795-1317 (1-1062)

QY 165 ThrGlyProValLysAlaGluGluProThrProThrProLysIleProLysLysArgGly 184
Db 141 TCAGGTGCTATGAACCCAGAA-----CCAAAAGCGTGGG 176

QY 185 ArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAspAsp 204
Db 177 AGAAGCGGAG---TTTGATCTTAACGTCGAGGATTCATCTGCGACGTGAAGATGAT 233

QY 205 ProGluAsnArgSerLysPheThrGluSerAlaSerAlaArgLysArgThrValThrAla 224
Db 234 GCTGACATGCGCTTAGTGCTATGAAGTGCTTCTGCCGAAAGAGAAACCGTGACAGCT 293

QY 225 GluGluArgGluArgAlaIleAlaAlaLysThrPheGluProThrAsnProPhePhe 244
Db 294 GAAGAAGAGAGAGAGCAATCAATGGCCCAAGCAATTTGAGCCTACTTAACCTTTCTGC 353

QY 245 ArgValValLeuArgProSerThrLeuThrArgGlyCysIleMetThrLeuProSerGly 264
Db 354 AGGTGCTCTTGGCACCATCATATCTGTACAGGGGATGCATTATGTACTTACCGTCTGTC 413

QY 265 PheAlaGluLysThrLeuSerGlyIleSerGlyPhe-IleLysValGluLeuAlaGlu-- 283
Db 414 TTTGCTCAGAAGCATCTTAAGTGGGGTTTCTGGATTCCATTAACTTCAGCTTCCTGATGG 473

QY 284 -LysGluTrpProValArgCysLeuThrLysAlaGlyArgAlaLysPheSerGlnGlyTr 303
Db 474 GAGACAGTGTGTGTACATGTCTGTATTAAGCAGGACGAGCTTAAGTTCACTCAGGAGATG 533

QY 303 pThrGluPheThrLeuGluAsnLeuGluGlyGluGlyAspValCysValPheGluLeuLe 323
Db 534 GTACGAATTACATTGGAGATAAATTTGGGAGAGGGGATGTCTGTCTTTCGAGCTGCT 593

QY 323 uArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGlu 339
Db 594 CAGATCGAGGGAATTCTGCTCAAGATTACCGTATTTTCGTGTAAAGGGA 642

RESULT 8
US-10-021-323-9015
; Sequence 9015, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill

US-10-088-187A-11 (1-341) x US-10-021-323-9015 (1-644)

QY 22 ValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeuThr 41
Db 1 ATCCCCATTAACCTTGTAGAAATTTGGGACGAACTTCTTGTGTCTGTCTCTCACT 60

QY 42 ValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPhe 61
Db 61 GTTCCTGATGTCATGTTTGGCGTGTAGGAATAAGGAAGCCGACAAACAAGTTTGGTTT 120

QY 62 GlnAspGlyTrpGlnGluPheValAspArgTrpSerIleArgIleGlyTrpLeuLeuIle 81
Db 121 CACGAGGTTTGGCAGGAATTTGTAGACGGTACCATATCCGTGTGGCTACTTACTTGGTT 180

QY 82 PheArgTrpGluGlyAsnSerAlaPheSerValTrpIlePheAsnLeuSerHisSerGlu 101
Db 181 TTCAGATATGAATGGAATTTCTTGTTCAGT----- 210

QY 102 IleAsnThrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla-Ar 121
Db 211 -----AAAACCGTA 219

QY 121 gLeuPheGluAspLeuGluAspGluAspAla-----GluValIlePhePr 136
Db 220 TCCATTGAGAATTTGAGATGACGATGATGATGATGATGATGATGATGATGATGATGAT 279

QY 136 oSerSerValTrpProSerProLeu-----ProGluSerThrValProAlaAsnLys 153
Db 280 TGGCTCTAAAGTTAATAACTGCACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 339

QY 153 sGlyThrAlaSerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluGluPr 173
Db 340 GGGCATGATGAATGCAACTCT-----TTTCGAGGTGCACTG----- 373

QY 173 oThrProThrProLysIleProLysArgGlyArgLysLysLysAsnAlaAspProGl 193
Db 374 -----CTACCCAAA---CCAAACAGCCTGGGAGGAAAAGCAGAGTTTGTACACGC 423

QY 193 uGluIleAsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTrpGlu 213
Db 424 TGAACCAAGATTCACTGTTGGACACGATATGATGATGATGATGATGATGATGATGATGAT 483

QY 213 uSerAlaSerAlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAl 233

Db 484 AAGTGTTCAGCGAAGAGAGAACTGTGCAACCGAAGAGAGAGAGCAATTAACGC 543
QY 233 aAlaIysThrPheGluProThrAsnProPhePheArgValValLeuArgProSerTyrLe 253
Db 544 AGCCAAATCGTTGAGCCATGATACCCCTTCTGCAGAGTTGCTTACGACCATCATATCT 603
QY 253 uTyrArgGlyCysIleMetTyrLeuProSer 263
Db 604 ATACCGGGGATGTATTATGTACTTACCATCA 634

RESULT 9

US-09-770-423-237/c

; Sequence 237, Application US/09770423

; Publication No. US20020040490A1

; GENERAL INFORMATION:

; APPLICANT: Goralach, Jörn

; APPLICANT: An, Yong-Qiang

; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.

; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang

; APPLICANT: Rameaka, Joshua G.

; APPLICANT: Page, Amy

; APPLICANT: Matthew, Abraham V.

; APPLICANT: Ledford, Brooke L.

; APPLICANT: Woessner, Jeffrey P.

; APPLICANT: Haas, William David

; APPLICANT: Garcia, Carlos A.

; APPLICANT: Krieker, Maja

; APPLICANT: Slader, Ted

; APPLICANT: Davis, Keith R.

; APPLICANT: Allen, Keith

; APPLICANT: Hoffman, Neil

; APPLICANT: Hurban, Patrick

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis

; FILE REFERENCE: 2028 (PARA-017PRV)

; CURRENT APPLICATION NUMBER: US/09/770,423

; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 60/178,512

; PRIOR FILING DATE: 2000-01-27

; NUMBER OF SEQ ID NOS: 999

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 237

; LENGTH: 426

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURES:

; NAME/KEY: misc.feature

; LOCATION: (1)..(426)

; OTHER INFORMATION: n = A,T,C or G

US-09-770-423-237

Alignment Scores:
Pred. No.: 2,87e-55 Length: 426
Score: 512.00 Matches: 96
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 2
Query Match: 28.51% Indels: 0
DB: 9 Gaps: 0

US-10-088-187A-11 (1-341) x US-09-770-423-237 (1-426)

QY 244 PheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeuProSer 263
Db 425 TTCAGAGTGTTCGACCATCTCTATATACAGAGTTGATCATGTATCTCTCTCT 366
QY 264 GlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGlu 283
Db 365 GGGTTTCTGAGAGTACTAAGTGGATCTCCGGTTTCATCAAGTCCAGNATGGAG 306
QY 284 LysGlnTrpProValArgCysLeuTyrIysAlaGlyArgAlaLysPheSerGlnGlyTrp 303

Db 305 AAACAATGGCTGTTCGATGTCTTACAAAAGCGGAGAGCCAAATTCAGTCAGGATGG 246
QY 304 TyrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPheGluLeu 323
Db 245 TAGCAATTCACCTCTAGAGAACCACTTAGGAGAGGAGACGCTCTGTGTGTTGAGCTGCTC 186
QY 324 ArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyrVal 341
Db 185 AGAACCCAGAGATTTGTTTTGAAAGTGACAGCCTTTCGAGTCACAGGATACGTC 132

RESULT 10

US-10-424-599-97591

; Sequence 97591, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovacic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 97591

; LENGTH: 628

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURES:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_59138C.1

US-10-424-599-97591

Alignment Scores:
Pred. No.: 1.37e-37 Length: 628
Score: 375.50 Matches: 74
Percent Similarity: 81.48% Conservative: 14
Best Local Similarity: 68.52% Mismatches: 19
Query Match: 20.91% Indels: 1
DB: 16 Gaps: 1

US-10-088-187A-11 (1-341) x US-10-424-599-97591 (1-628)

QY 182 LysArgGlyArgLysLysAsnAlaAspProGluGluIleAsnSerAlaProArg 201
Db 303 AAAACCGCAAAAAGAGCGGAATCAGAACCTATGGCGAGAACCTTCTGGTGAAT 362
QY 202 AspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThr 221
Db 363 GAAGAGGAGCGAGAAATCGCTATAGGTTTATGAAAGTGCATCTGCAAGGAACGACC 422
QY 222 ValThrAlaGluGluArgGluArgAlaIleAsnAlaLysThrPheGluProThrAsn 241
Db 423 GTGACGGCAGAGAAAGAGAAAGGCAATTAATGCAATCAAAAACATTTGAACCGACTAT 482
QY 242 ProPheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeu 261
Db 483 CTTTCTCCGAGTTGCTCTCGACCCCTCTATTATTATAGGGATGATATATATATG 542
QY 262 ProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeu 281
Db 543 CTTTCCACCTTTGCTGAAAGAAATTTGAATGGGTTTCGGGATTCATCAAACTTCAGCTG 602
QY 282 AlaGluLysGlnTrpProVal 288
Db 603 TCCAATGGTAGACAGTCGTCGTT 626

RESULT 11

US-10-767-795-1320

; Sequence 1320, Application US/10767795

; Publication No. US2004018130A1

; GENERAL INFORMATION:

APPLICANT: Kovalic, David K.
APPLICANT: Cao, Yongwei
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(53534)B
CURRENT APPLICATION NUMBER: US/10/767,795
CURRENT FILING DATE: 2004-01-30
NUMBER OF SEQ ID NOS: 117596
SEQ ID NO 1320
SEQ ID NO 1320
LENGTH: 650
TYPE: DNA
ORGANISM: Gossypium hirsutum
FEATURE:
OTHER INFORMATION: Clone ID: LIB3543-054-Q1-K6-E2
US-10-767-795-1320
Alignment Scores:
Pred. No.: 7,98e-32 Length: 650
Score: 330.50 Matches: 70
Percent Similarity: 54.82% Conservative: 21
Best Local Similarity: 42.17% Mismatches: 20
Query Match: 18.40% Indels: 56
DB: 3 Gaps: 3
US-10-088-187A-11 (1-341) x US-10-767-795-1320 (1-650)
QY 1 MetProArgProPhePheHisLysLeuLeuPheSerThrThrLeuGlnLysArgLeu 20
DB 201 ATGCCACGCCCTTTTTCATTAAGCTTATCTCTACCCCTCCAAAGAAAG----- 254
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
DB 254 ----- 254
QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
DB 255 -----AAACT-GGAAATAAGAAAGTTGACAAACAAGTTGG 289
QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTrpSerIleArgIleGlyTrpLeuLeu 80
DB 290 TTTTCAGGAAGTTGGCAGGAGTTTCTAGACGCTTACTATATTCGTGTGGCTACACTG 349
QY 81 IlePheArgTrpGluGlyAsnSerAlaPheSerValTrpIlePheAsnLeuSerHisSer 100
DB 350 GTTTTCAGATACCAAGAAATCTGCTTCAGTGTTAGTATATTAATTTGTATACCTCG 409
QY 101 GluIleAsnTrpHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
DB 410 GAATAAAGTATCAGACTAATGCTTGGTGGTACTCAATCAATCAGCGAAACAATAT 469
QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTrp 140
DB 470 CCG---TTTGAACAACCTTGAAGATGATGAATGC----- 499
QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTrpAlaSerSerAlaIle 160
DB 500 -----ATCTCTCCAGCACTT 514
QY 161 GlnThrLeuPheThrGly 166
DB 515 CAGAAATTTGTTGGTGGG 532
RESULT 12
US-10-739-930-3743/c
Sequence 3743, Application US/10739930
Publication No. US20040216190A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-21(53377)B
CURRENT APPLICATION NUMBER: US/10/739,930

CURRENT FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 3743
LENGTH: 2207
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(2207)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: GLYMA-23APR03-CLUSTER63260_1
US-10-739-930-3743
Alignment Scores:
Pred. No.: 8.49e-31 Length: 2207
Score: 329.00 Matches: 94
Percent Similarity: 43.77% Conservative: 57
Best Local Similarity: 27.25% Mismatches: 130
Query Match: 18.32% Indels: 64
DB: 10 Gaps: 10
US-10-088-187A-11 (1-341) x US-10-739-930-3743 (1-2207)
QY 6 PheHisLysLeuIlePheSerThrThrIleGlnGlyArgLeuArgValProAspLys 25
DB 1757 TTTTTCAGATCATACCGCTCACAATGTTTCATGAGAAAGCTAATGATGCCCAATAAG 1698
QY 26 PheValSerLysPheLysAspGluLeuSerValAlaValAlaLeuThrValProAspGly 45
DB 1697 TTTGTGAAGAAATATGAAACCGCTTACAAAATCTCTTTCTGAGACTCCAAATGGT 1638
QY 46 HisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGlyTrp 65
DB 1637 GCTCAATGGAAATGATTTTGAAGAACGTCGCTGAAATATGTTTCAAAAGGATGG 1578
QY 66 GlnGluPheValAspArgTrpSerIleArgIleGlyTrpLeuLeuIlePheArgTrpGlu 85
DB 1577 AAAGAGTTTCAGAGATATCTCTCTAGCTCATGGCCATCTTTTGGTTTTCAGATGGAT 1518
QY 86 GlyAsnSerAlaPheSerValTrpIlePheAsnLeuSerHisSerGluIleAsnTrpHis 105
DB 1517 GTAACTTCTCATTTTCAGGTACACATCTTTGATTTGATGCTTCCCTTAGAGATTGAGTAC--- 1461
QY 106 SerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArgLeuPheGluAsp 125
DB 1461 ----- 1461
QY 126 LeuGluAspGluAspAlaGluValIlePheProSerSerValTrpProSerProLeuPro 145
DB 1460 -----CTTACCGAAATA-----ATAAAA 1443
QY 146 GluSerThrValProAlaAsnLysGlyTrpAlaSerSerAlaIleGlnThrLeuPheThr 165
DB 1442 GGTAAACAGCTCCATAGAAAGAAATGAATCTCCAGGTGATGAACATTTGGAGTGT 1383
QY 166 GlyProValLysAlaGluGluProThrProThrProLysIleProLysArgGlyArg 185
DB 1382 -----CACAGATCAGGTCAA 1368
QY 186 LysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAspAsp----- 203
DB 1367 AAGAAAGAAATTAAGTCACTGATTTCTTCAACAAATGTCAAATGAGAGTCGCAAGTGT 1308
QY 204 ---AspProGluAsn-----ArgSerLysPheTrpGluSerAlaSerAla 217
DB 1307 GTTAAAGTTGAGACACGATCATCTGCCAAGCAGGCTCTGCTACACATGCCCAAAA 1248
QY 218 ArgLys---ArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaLysThr 236
DB 1247 TGCAAAAGAAATTCGAAGCTATGATTAACCAAGTTACAGCTCTTGTATGAGCAAGTCT 1188
QY 237 PheGluProThrAsnProPhePheArgValValLeuLeuArgProSerTrpLeuTrpArgGly 256

```

Db      1187 TTCAAAATCTGTGAATCCATCTCTTTGACCGTCATCGAACCCACATCAGTTCTCTAC 1128
Qy      257 CysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSer-----GlyIleSer 274
Db      1127 GGTAGTCTGACATTAACCAATGAAGTTCTGTAGGAGTCACCTTCATTGTGCAAGAAGCGG 1068
Qy      275 GlyPheIleLysValGln---LeuAlaGluLysGlnTTPProValArgCys---LeuTyr 292
Db      1067 AGACTTATCAGCCTTCAGGCTTGAAGTGGAGAAATTGGCTCGCAAAAGTATCAGATCCAT 1008
Qy      293 LysAlaGlyArgAla---LysPheSerGlnGlyTyrTyrGluPheThrLeuGluAsnAsn 311
Db      1007 AAACAAAGACTGCTATAGATTAAACTCTCATGGAACGCAATTGTCTAAGGACAATAAC 948
Qy      312 LeuGlyGluGlyAspValCysValPheGluLeuLeuArgThrArgAspPheValLeuLys 331
Db      947 TTGAAAGTTGGTGGATGTTTGCATCTTTGAACCTCGTTCACTGGAACAACTAACCTTCCTG 888
Qy      332 ValThrAlaPheArg 336
Db      887 GTTCACATCTTCAGA 873

```

RESULT 13

US-10-425-115-89685

; Sequence 89685, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 89685

; LENGTH: 2338

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MR74577_181795C.1

US-10-425-115-89685

Alignment Scores:

```

Pred. No.:      1,93e-27      Length:      2338
Score:           303.00      Matches:      109
Percent Similarity: 40.64%      Conservative: 56
Best Local Similarity: 26.85%      Mismatches: 155
Query Match:      16.87%      Indels:      86
DB:              18          Gaps:       17

```

US-10-088-187A-11 (1-341) x US-10-425-115-89685 (1-2338)

```

Qy      4 PropPheHisLeuLeuPheSerThrIleGlnGluLysArgLeuArgValPro 23
Db      761 CGCAGTCTTCAAGGTCTTCTCCCGAGCAGTCACCGAG---CGGCTGAGATCCCA 817
Qy      24 AspLysPheValSerLysPheLysAspGluLeuSerVal---AlaValAlaLeuThrVal 42
Db      818 ACCATGTTCAACAGCACCTCAAGGAGCAGCAGCCAAACCGGAGCAGTTTCCCTGAGAGGC 877
Qy      43 ProAspGlyHisValTyrArgValGlyLeu-----ArgLysAlaAspAsnLysIleTyr 60
Db      878 CCGAGCGGCAACAGATGGCAGCAGCGCTCGTTCGAGTCCGAGTCGGAGGCCCGCGTGG 937
Qy      61 ---PheGlnAspGlyTyrGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeu 79
Db      938 TCGTTCGACCGAGGCTGGAGAGATTGTTCAGGACCACCTCCCTCAGGCTAGGCCACTTC 997
Qy      80 LeuIlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHis 99

```

RESULT 14

US-10-437-963-89631/c

; Sequence 89631, Application US/10437963

; Publication No. US2004012343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

```

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 89631
; LENGTH: 1477
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_8836C.1
; US-10-437-963-89631

Alignment Scores:
Pred. No.: 66-25 Length: 1477
Score: 281.00 Matches: 104
Percent Similarity: 42.93% Conservative: 60
Best Local Similarity: 27.23% Mismatches: 152
Query Match: 15.65% Indels: 66
DB: 17 Gaps: 15

US-10-088-187A-11 (1-341) x US-10-437-963-89631 (1-1477)
QY 3 ArgProPheHisLeuLeuPheSerSerThrLeuGlu---LysArgLeuArg 21
DB 1447 AGAGCTTCTTCGGTGTCTTC---ACCTGCATCTCTCGAGCGGATGAAA 1397
QY 22 ValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeuThr 41
DB 1396 ATCCCATCATCATCAACCAATGCTTCAGAAATCAACCTACTGCGATGCTTCTCGTT 1337
QY 42 ValProAspGlyHisValTrpValGlyLeuArgLysAlaAspAsnLysIleTrpPhe 61
DB 1336 GATCGAAGTGGAAATAAATGAGTGCAGAGCTGACTTCAGACTCTCGAGGATTCCTTT 1277
QY 62 GlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTrpLeuLeuIle 81
DB 1276 GTGCATGTAAGGAAAGAAATTTGTCAGAGATACTCAATCACTGCTGGTCAATTTAGTA 1217
QY 82 PheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSerGlu 101
DB 1216 TTCACATTATGATAAACGGTCAACAATCTCAGTGACAGTCTTT---GAACCTTCAGGC 1163
QY 102 IleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
DB 1162 ATTGACAAGATATCAACT-----TTCTCTGCTCACCT-----TCCAAAATGTG 1118
QY 122 LeuPheGluAspLeuGluAspGluAlaGluValIlePheProSerSerValTyrPro 141
DB 1117 ATCATTAACACAAAGTAGTATGAGGGGCGATGGTCCCGCAGCTATCACACA----- 1064
QY 142 SerProLeuProGluSerThrValProAlaAsnLys-----GlyTyr 155
DB 1063 -----GAGAAAATGGTCCCGCCCTGAGGAAAAATAATGAATCACAGGAAAG 1016
QY 156 AlaSerSerAlaIleGlnThrLeuPheThrGlyProVal-----LysAlaGlu 171
DB 1015 AGAACTAGAGATAGATAGATACCTATGGAAGATAGGTTGTTGTTATTAAGAAATCATCT 956
QY 172 GluProThrProThrProLysIleProLysArg-----GlyArgLys 186
DB 955 GAGCTTAATGTTGCGAAAGTTCTAGAGAAAACGTGCAGAGAGCTTCTGCTGCAAGTCC 896
QY 187 Lys-----LysAsnAlaAspProGluLeuIle 195

```

```

DB 895 AAGTAACATCGACAAAGCCATAACAGTACCAGAGGTTTCATCATGCGAGCTCAGATGAAGAT 836
QY 196 AsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGluSerAla 215
DB 835 AATTCCTTAGC---AAGTCACCAATCCACCATTCCTGATGCGCTTCTGCTGCTGAG 779
QY 216 SerAlaArg----- 218
DB 778 GTGTCCTCGGCTGTAGATGTTTCAAAGGGTCAAAGACAACTAACAGTAATATCACAG 719
QY 219 LysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGlu 238
DB 718 AGCGCTCCGGTAACTGAAGCAGAGAGGACCATGCTCTTCAAAGGCGAAGAGATTAAA 659
QY 239 ProThrAsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIle 258
DB 658 TCTAAGAATCTCTTTCGCGTGCAAAATAATGATGAGTCAATATGCTATGTTGGATTTC 599
QY 259 MetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLys 278
DB 598 ATGAATATCCCATGTGATTTTCCGTCGAGTGTCTTCCCATACCAACAAAGGATCACA 539
QY 279 ValGlnLeuAlaGlu---LysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAla 297
DB 538 CTCTGGGATCCACAGGAGGCTTGGAAAGTTAACTATGCTACTACAGTATCGTCT 479
QY 298 -----LysPheSerGlnGlyTyrTyrGluPheThrLeuGluAsnLeuGlyGluGly 315
DB 478 GTTGGTCTTTCAGTGTGCTGGGCAAAATTTGCTGTGGCAACATCTGGAGAGTTT 419
QY 316 AspValCysValPheGluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPhe 335
DB 418 GATGTGTGTCTCTCGAGCTTGTCCAGAGGATAAC-----ATAAAGTGCACATCTAC 365
QY 336 ArgVal 337
DB 364 AGAGTT 359

RESULT 15
US-10-425-114-32044
; Sequence 32044, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32044
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73255A02_FLI
; US-10-425-114-32044

Alignment Scores:
Pred. No.: 7,86e-25 Length: 1455
Score: 280.00 Matches: 89
Percent Similarity: 44.55% Conservative: 54
Best Local Similarity: 27.73% Mismatches: 136
Query Match: 15.59% Indels: 42
DB: 16 Gaps: 13

US-10-088-187A-11 (1-341) x US-10-425-114-32044 (1-1455)

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 30, 2004, 06:15:53 ; Search time 6749 Seconds
(without alignments)
1841.154 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796
Sequence: 1 MRPFPFKLIFSTIQEKL.....LLRTDFVLKVTAFRNVYV 341

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-O=/csp2_1/USPTO.spool/US10088187/runat_29122004_115842_127/app.query.fasta_1.519
-DB=EST_QFMT=fastap SUPFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.coi -LIST=45
-DOALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10088187 @CNC 1 1 3437 @runat_29122004_115842_127 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: gb_est1:
2: gb_est2:
3: gb_hcc:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gss1:
9: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1789	99.6	1307	CNS0A70B	BX822585 Arabidops
2	1769	98.5	1441	CNS0A6ON	BX823300 Arabidops
3	1717	95.6	1375	CNS0A6SO	BX824523 Arabidops
4	1142.5	63.6	1346	CNS0A6P	BX816536 Arabidops
5	1141.5	63.6	1420	CNS0A03Z	BX816391 Arabidops
6	1141.5	63.6	1490	CNS0A086	BX815859 Arabidops
7	1061.5	59.1	880	CNS08969	BX841843 Arabidops
8	1016.5	56.6	780	CK120978	CK120978 204614.p1
9	934.5	52.0	906	CO108220	CO108220 GR_Eb003

10	932	51.9	701	6	CD822418	CD822418 BN25.045C
11	860	47.9	1041	5	BX835682	BX835682 BX835682
12	811	45.2	834	7	CO108208	CO108208 GR_Eb003
13	789.5	44.0	669	1	AV821829	AV821829 AV821829
14	786	43.8	645	6	CD823066	CD823066 BN25.0471
15	771.5	43.0	834	7	CO094348	CO094348 GR_Ea16G
16	707.5	39.4	612	6	CD813712	CD813712 BN15.020L
17	682	38.0	589	1	AV540912	AV540912 AV540912
18	667	37.1	591	1	AI996485	AI996485 701666887
19	660	36.7	785	8	BZ061379	BZ061379 lki10h12.
20	657	36.6	614	1	AV539303	AV539303 AV539303
21	655.5	36.5	748	7	CO084070	CO084070 GR_Ea49H
22	652	36.3	841	7	CO132469	CO132469 GR_Eb45F
23	650	36.2	721	8	BH969696	BH969696 ocd94e04.
24	626.5	34.9	699	8	BZ464608	BZ464608 BONRX74TR
25	623.5	34.7	870	4	BM360659	BM360659 GA_Ea003
26	623	34.7	872	7	CO128357	CO128357 GR_Eb25B
27	620.5	34.5	591	6	CA781580	CA781580 024G12AF
28	620	34.5	809	7	CO099865	CO099865 GR_Ea24P
29	617	34.4	617	4	BM814806	BM814806 ESTS92900
30	603	33.6	631	1	AJ504308	AJ504308 AJ504308
31	597.5	33.3	712	7	CO132468	CO132468 GR_Eb45F
32	583.5	32.5	624	2	AW584452	AW584452 N210502e
33	579	32.2	636	6	CD038665	CD038665 UTPPT010
34	577.5	32.2	821	7	CO108207	CO108207 GR_Eb003
35	577.5	32.2	822	7	CO108219	CO108219 GR_Eb003
36	576	32.1	726	5	BX835627	BX835627 BX835627
37	575	32.0	453	6	AW705298	AW705298 8k59c02.Y
38	564.5	31.4	754	6	CA990755	CA990755 EST644263
39	556.5	31.0	732	6	CD835772	CD835772 BN45.046J
40	553.5	30.8	537	1	BM143791	BM143791 saj48910.
41	543	30.2	317	1	AV526843	AV526843 AV526843
42	542	30.2	471	1	AI992821	AI992821 701493711
43	535.5	29.8	448	7	CN493904	CN493904 Mdfw2004a
44	534.5	29.8	480	1	AL383040	AL383040 MCBCL1D09
45	533.5	29.7	812	4	BM357983	BM357983 GH_Ea000

ALIGNMENTS

RESULT 1
CNS0A70B 1307 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTFFB50ZB05 of flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
ACCESSION BX822585
VERSION BX822585.1 GI:42465399
KEYWORDS HTC; GSLT_cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1307)
AUTHORS Castellani, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,
Menard, G., Cruaud, C., Quetier, P., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1307)
AUTHORS Genoscope.
JOURNAL Direct Submission
TITLE Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
[http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length)
<http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis>.

FEATURES

Location/Qualifiers

source

1..1307
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLFB50ZE05"
/issue_type="Flowers and buds"
/plasmid="pCMVSPORT 6"
complement(1..1307)
/gene="At3g18990"

ORIGIN

Alignment Scores:

Pred. No.: 1.53e-181 Length: 1307
Score: 1789.00 Matches: 340
Percent Similarity: 99.71% Conservative: 0
Best Local Similarity: 99.71% Mismatches: 1
Query Match: 99.61% Indels: 0
DB: 3 Gaps: 0

US-10-088-187A-11 (1-341) x CNS0A70B (1-1307)

QY 1 MetProArgProPhePheHisLysLeuLeuPheSerSerThrIleGlnLysArgLeu 20
DB 75 ATGCCACGCCCTTCTTCATAGTTGATTTCTCATCTATTCACATATTCAGAAAAAGCTTG 134
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
DB 135 AGGGTCCACAGATAAGTTGTGAGTAAATTCAGAGTATGAGCTTTCGGTGTCTGTCATCTC 194
QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
DB 195 ACAGTACCTGATGGTGTATGTTGGCGGTGAGGACTAAGGAAGCTGACAAACAAATTTGG 254
QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80
DB 255 TTTCAAGATGGTTGGCAAGAGTTGTTCACCTTACTCCATTCGATGGTATCTTTTG 314
QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
DB 315 ATTTTATAGATATGAAGGAACCTCTGCTTACGGTCTACATTTTCAATTTATCCACTCT 374
QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
DB 375 GAGATCAATTTACCATTCACCGGTCTCATGGATTCGGTCAACACCTTCAACGCCGCC 434
QY 121 ArgLeuPheGluAspLeuGluAspGluLeuPheAlaGluValIlePheProSerSerValTyr 140
DB 435 CGTTTGTGAGACCTTGAAGATGAGATGCGGAGTCACTTCTCTCTCTCTGTTGATC 494
QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
DB 495 CCATCACCATTCTCTGAGTCTACAGTACGACCAACAAAGGATAGTGTAGTTCAGCCATC 554
QY 161 GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro 180
DB 555 CAAACCTTGTTCCTGGACAGTCTAAGCTGAAGCAACGCCAACCCCAAAATACCT 614
QY 181 LysLysArgGlyArgLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro 200
DB 615 AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 674
QY 201 ArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220

DB 675 CGAGATGATGATCCAGAGAACCGTTCAAAAGTTCTACGAGAGTGCTTCTCGCAGAAAGAGA 734
QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAlaIleAlaLysThrPheGluProThr 240
DB 735 ACCGTGACTCGAGAGAAAGAGAGAGAGCCATCAATGACGACCAACCGTTGGAACCAACA 794
QY 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260
DB 795 AACCTTTCTTCAGAGTGTTCTCGACCATCTCTATACAGAGGTTCATCATGTAT 854
QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280
DB 855 CTTCCTCTCGGTTTCTGCTGAGAGTACCTAAGTGGATCTCCGGTTTCATCAAGTCCAG 914
QY 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300
DB 915 CTTGCGGAGAAACAATGGCCTGTTCGATGTCTCTACAAAGCCGGGAGAGCCAAATTCAGT 974
QY 301 GlnGlyTrpTyrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPhe 320
DB 975 CAAGATGTAGCAATTCATCTAGAGAACACTTAGGAGAGAGAGAGCTCTGTGTGTTT 1034
QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340
DB 1035 GAGCTGCCAGAACCCAGAGATTTCTGTTTGAAGTGACAGCTTTTCGAGTCAACGAGTAC 1094
QY 341 Val 341
DB 1095 GTC 1097
RESULT 2
CNS0A6ON
LOCUS
DEFINITION
CNS0A6ON 1441 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSUTLS202G12 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION
VERSION BX823300.1 GI:42464860
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM
REFERENCE
AUTHORS Castell, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1441)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
Web : www.genoscope.cns.fr)
COMMENT
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castell
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
[http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length)
<http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis>.
Location/Qualifiers
1..1441
FEATURES
source

/organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLTSL20ZG12"
 /tissue_type="Adult vegetative tissue"
 /plasmid="pCMVSPORT 6"
 complement(1..1441)
 /gene="At3g18990"

gene

ORIGIN

Alignment Scores:

Pred. No.: 2,518-179 Length: 1441
 Score: 1769.00 Matches: 340
 Percent Similarity: 99.71% Conservative: 0
 Best Local Similarity: 99.71% Mismatches: 1
 Query Match: 98.50% Indels: 1
 Ds: 3 Gaps: 0

US-10-088-187A-11 (1-341) x CNS0A60N (1-1441)

QY 1 MetProArgProPheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20
 Db ATGCCACGCCCTTCTTCATTAAGTTGATTTCTCATCCATATCCAGAGAAAGCTGTG 291
 QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
 Db AGGGTCCAGATAAAGTTGTGAGTAATAATCAAGGATGAGCTTCGGTGTGCTGCTGCACCTC 351
 QY 41 ThrValProAspGlyHisValTrpArgValGluArgLysAlaAspAsnLysIleTrp 60
 Db ACAGTACCTGATGGTCATGTTTGGCGTGTAGACATCAAGAAAGCTGACAAACAAATTTGG 411
 QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80
 Db TTTCAAGATGGTTCGACAGAGTTTGTGACCGTACTTCATTCGATTCGATTCGATTCGATTCG 471
 QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
 Db ATTTTATAGATATGAAGAAACTCTGCCCTTCAGCGTCACTATTTCAATTTATCCACTCT 531
 QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
 Db GAGATCAATTAACCATTCACCGCTCTCATGGATTCGCTCAACACCATTCACCAACGCCGCC 591
 QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140
 Db CTTTGTGTTGAAGACTTGAAGATGAGATGCCGAGGTCATCTTCTCTCTCTCTCTCTCTCTCT 651
 QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
 Db CCATCACCACCTTCCTGAGTCTACAGTACCAGCCCAACAAAGGTATGCTAGTTCAGCCATC 711
 QY 161 GlnThrLeuPheThrGlyProValLysAlaGluLuproThrProThrProLysIlePro 180
 Db CAACCTTTGTTTCACTGGACCACTTAAAGCTGAAGAGCCCAACGCCCAACCCCAAAATACCT 771
 QY 181 LysLysArgGlyValGlyLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro 200
 Db AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 830
 QY 201 ArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220
 Db CGAGATGATGATCCAGAGACCGTTCAAGTTCTACGAGAGTCTTCGCGAGAAAGAGA 890
 QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240
 Db ACCGTGACTGCAGAGAAAG 950
 QY 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260
 Db AACCTTCTTCTCAGAGTGGTCTCGGACCATCTCTATCTATATACAGAGGTTGTCATCATGTAT 1010

QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280
 Db CTTCTCTCTGGGTTTGCTGAGAAAGTACCTTAAGTGGGATCTCGGGTTTCATCAAGTCCAG 1070
 QY 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGluArgAlaLysPheSer 300
 Db CTTGCGGAGAAACATGGCTGTTTCGATGTTCTCTACAAAGCCGAGAGACCAATTCAGT 1130
 QY 301 GlnGlyTyrTrpGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPhe 320
 Db CAAGGATGTTACGAATTCACCTAGAGAACTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1190
 QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValLeuGluTyr 340
 Db GAGCTGCTCAGAACCCAGAGATTCGTTTGAAGTGACAGCCCTTTTCGAGTCAACGAGTAC 1250
 QY 341 Val 341
 Db 1251 GTC 1253

RESULT 3
 CNS0A60S
 LOCUS
 DEFINITION
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSUPGH542B04 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1375)
 Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
 Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
 Unpublished
 2 (bases 1 to 1375)
 Genoscope.
 Direct Submission
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
 URV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length
 http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
 location/Qualifiers
 1. 1375
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLTGH542B04"
 /tissue_type="Hormone Treated Callus"
 /plasmid="pCMVSPORT 6"
 complement(1..1375)
 /gene="At3g18990"

FEATURES
 source
 gene
 ORIGIN


```
QY 99 sSerGluIleAsnTyrHisSerThrGly-----LeuMetAspSerAlaHisAs 115
Db 351 CTCTGAGATCAACCACTTCTAGTAGTGAAGCTCTTATGCAAAATGGATTCCGCAGAA 410
QY 115 nHisPhe---LysArgAlaArgLysPheGluAspLeuGluAspGluValI 134
Db 411 TCAGTTCAACAACGCGTCTGATTTGGAAGATCTCTGAACCTCAAGATGCGAAGTCA 470
QY 134 ePheProSerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysG 154
Db 471 TTATCATCGAAC-----CCTGAATCTACTGAACCACTGAATAAGG 512
QY 154 YTYR---AlaSerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluPr 173
Db 513 TTATGCGCGTCTTACAGCCATCAAAAGCTTTTTCAAA---GAATCTAAAGCTGAAGAA-- 567
QY 173 oThrProThrProLysIleProLysArgLysGlyArgLysLysAsnAlaAspProG 193
Db 568 -----ACGCCAAGGTACTTAAGAAAGAGAGAGAAAGAAAGAAATCCTAATCCCGA 620
QY 193 uGluIleAsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrG 213
Db 621 GGAAGTAAACTCTTCAACTCCCGTGGAGATGACTCAGAGAACCGCTCAAGTCTACGA 680
QY 213 uSerAlaSerAlaArgLysArgThrValThrAlaGluGluArgGluArgAlaAsnAl 233
Db 681 GAGTGTCTCTGCTAGAAAGAGAACTGTAATCTCAGAGGAAAGAGAGAGCGCTCAATGC 740
QY 233 alalysThrPheGluProThrAsnProPheArgValValLeuArgProSerTyrI 253
Db 741 AGCCAAACATTCGAACCAACAACTCTTACTTTAGAGTTGTCTGCGACCATCATCT 800
QY 253 uTYrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGly 273
Db 801 ATACAGAGGTTCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 860
QY 273 eSerGlyPheIleLysValGlnLeuAlaGluLysGlnTrpProValArgCysLeuTyr 293
Db 861 ATCTGCTTTCATCAAGCTCCAGCTCGGTGAGAAACAAATGCGCAGTGAGTGCCTTACAA 920
QY 293 salaglyArgAlaLysPheSerGlnGlyTyrTrpTyrGluPheThrLeuGluAsnLeu 313
Db 921 AGCAGGAGAGTAAAGTTAGCCAAAGGATGGTATGAGTTTCACTGAGTCAACGAGAAATAGG 980
QY 313 yGluGlyAspValCysValPheGluLeuLeuArgThrArgAspPheValLeuLysVal 333
Db 981 CGAAGAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1040
QY 333 rAlaPheArgValAsnGluTyrVal 341
Db 1041 CGCCTTCGTGTCATGAGTATGTG 1065

RESULT 7
CNS09Y69
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSL7LS76ZC03 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
BX841843
CNS09Y69
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 880)
Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
```

```
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_BF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
Location/Qualifiers
1..880
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strains="Col-0"
/db_xref="taxon:3702"
/clone="GSL7LS76ZC03"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
complement(1..880)
/gene="At1g49480"
source
gene
ORIGIN
Alignment Scores:
Pred. No.: 1.9e-103 Length: 880
Score: 1061.50 Matches: 226
Percent Similarity: 83.67% Conservative: 25
Best Local Similarity: 75.33% Mismatches: 22
Query Match: 59.10% Indels: 28
DB: 3 Gaps: 9
US-10-088-187A-11 (1-341) x CNS09Y69 (1-880)
QY 51 GlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGlyTrpGlnGluPheValAsp 70
Db 1 GCAGTAAAGAAAGCCACACAAATCTGTTTCAAGACGGTTGCGAGGAGTTTGTCAAC 60
QY 71 ArgTyrSerIleArgIleGlyTyrLeuLeuIlePheArgTyrGluGlyAsnSerAlaPhe 90
Db 61 CATTCTCCATTCGGATTGGT-----TTCAGATACAAA-----GTT 96
QY 91 SerValTyrIlePheAsnLeu---Ser-HisSerGluIleAsnTyrHisSerThrGly-- 108
Db 97 ACAGTCTACATTTTCAATTTATCTCTCCACACTCTGAGATCAACACCACTTCTAGTAGTGA 156
QY 109 -----LeuMetAspSerAlaHisAsnHisPhe---LysArgAlaArgLeuPheG 124
Db 157 AGCTCTTATGCAATGGATTCCGACAGATCAGTTCACAAACGCTGCTGATTTGTGA 216
QY 124 uAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyrProSerProbe 144
Db 217 AGATCCTGAACCTCAAGATGCTTAAGTTCATTTATCCTCAAC----- 259
QY 144 uProGluSerThrValProAlaAsnLysGlyTyr---AlaSerSerAlaIleGlnThrI 163
Db 260 -CCTGAATCTACTGAACCAAGTGAATGAGGTTATGCGGTTTCTACAGCCATCCAAAGCT 318
QY 163 uPheThrGlyProValLysAlaGluGluProThrProThrProLysIleProLysIle 183
Db 319 TTTTCAA---GAATCTAAAGCTGAAGAG-----ACGCCAAGGTACTTTAAGAGAG 366
QY 183 gGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAspAs 203
```


Qy	306	eThrLeuGluaaAsnLeuGlyGLuGLyAaspValCyValpHeGluLeuTeu	323
Db	728	CACACTCGAGAACAAATATAGCGCAAGGAGATGTATGTGTGTTTCAGCTACTC	779
RESULT 9			
CO108220			
LOCUS			
DEFINITION	CO108220	GR_EB0039N17.r GR_Eb Gossypium raimondii cDNA clone GR_EB0039N17	linear EST 16-JUN-2004
ACCESSION		3', mRNA sequence.	
VERSION		CO108220	
KEYWORDS		EST.	
SOURCE		CO108220.1 GI:48806906	
ORGANISM		Gossypium raimondii	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 906)	
AUTHORS		Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C., Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and Wing,R.A.	
TITLE		Global assembly of Cotton ESTs	
JOURNAL		Unpublished (2004)	
COMMENT		Contact: Rod A. Wing Arizona Genomics Institute The University of Arizona Forbes Building Room 303, Tucson, AZ, 85721-0036, USA Tel: 520 626 9595 Fax: 520 621 1259 Email: http://genome.arizona.edu Plate: 0039 row: N column: 17.	
FEATURES		Location/Qualifiers	
source		1..906	
		/organism="Gossypium raimondii"	
		/mol_type="mRNA"	
		/db_xref="taxon:297730"	
		/clone="GR_EB0039N17"	
		/tissue_type="floral"	
		/dev_stage=" 3 to +3 DPA"	
		/lab_host="DH10B"	
		/clone_lib="GR_Eb"	
		/notes="Vector: pCMV.SPORT-6.1; Site.1: NotI; Site.2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."	
ORIGIN			
Alignment Scores:			
Pred. No.:	9..39e-90	Length:	906
Score:	934.50	Matches:	189
Percent Similarity:	75.24%	Conservative:	45
Best Local Similarity:	60.77%	Mismatches:	54
Query Match:	52.03%	Indels:	23
DB:	7	Gaps:	7
US-10-088-187A-11 (1-341) x CO108220 (1-906)			
Qy	9	LeuilePheSerSerThrIleGlnGluIyAArgLeuArgValProAptysPheValSer	28
Db	2	CTTATTCTCTCTACCACTCCCAAGAAAGAACTGAGGATCCCTGATACTTCGTTAAG	61
Qy	29	LyvPheIyAaspGluLeuSerAlaIaValAlaLeuThrValProAptGlyHisValTrp	48
Db	62	AAATTACGGATGAACATTTCTGTGTCTGCGCTCTCACTTCTCAGCGTCAGTTTGG	121
Qy	49	ArgValGlyLeuArgIySAlaAaspAsnIySileTrpPheGlnAaspGlyTrpGlnGluPhe	68
Db	122	CGTGTAGGATAAAGAAAGTTGACACAAAGTTTGGTTTCAGGAAAGTTGCGCAGAGTTT	181
Qy	69	ValasPargTyrSerIleArgIleGlyTyrLeuLeullePheArgTyrGluGlyAasnSer	88
Db	182	CTAGAGGTTTACTATATTCTGTGTGGTACTGCTATTTGTTTATAGATACGAAGAAATCT	241

and <http://genoplatte-info.infobiogen.fr/>.

```

FEATURES
  source
    1..701
      /organism="Brassica napus"
      /mol_type="mRNA"
      /cultivar="Jet Neuf"
      /db_xref="taxon:3708"
      /clone="BN25045C06"
      /issue_type="seed"
      /clone_lib="BN25"

ORIGIN
  Alignment Scores:
    Pred. No.: 1,21e-89 Length: 701
    Score: 932.00 Matches: 180
    Percent Similarity: 92.12% Conservative: 7
    Best Local Similarity: 88.67% Mismatches: 14
    Query Match: 51.89% Indels: 2
    DB: 6 Gaps: 2

US-10-088-187A-11 (1-341) x CD822418 (1-701)

QY 1 MetProArgProPhePheHisLysLeuLeuPheSerSerThrIleGlnGluLysArgLeu 20
DB 92 ATGCCACGCCCTTTCTCCACAGCTGATTTCTCATCCACTATCCAGAAAAACGTCGT 151
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
DB 152 AGAGTTCCTGTAAGTTGTGAGTAATTCAGGACGAGCATCCGTTGCGGTGCTCTC 211
QY 41 ThrValProAspGlyHisValTyrArgValGlyLeuArgLysAlaAsp---AsnLysIle 59
DB 212 ACAGTACCTGATGTCATGTTTGGGTGTAGGACTAAGGAAAGCTGACACACAAAATC 271
QY 60 TrpPheGlnAspGlyTyrGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeu 79
DB 272 TGGTTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCAATCCGCAATTGGTTACCTT 331
QY 80 LeuIlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHis 99
DB 332 TTGATATTAGATACGAAGGCAACTCTGCTTCACGGTCTACATTTCACACTTATCACAC 391
QY 100 SerGluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArg 119
DB 392 TCCGGCATCACTACCACTCCACCGGTCTCATGGATTCTGCACACACCACTTCAAGCT 451
QY 120 AlaArgLeuPheGluAspLeuGluAspGluAlaGluValIlePheProSerSerVal 139
DB 452 GCGCGTTTGTGTAAGACCCCGAAGATGTTGAGGTTGTGAGGTTGTTCACCCCTCTCTTTG 511
QY 140 TyrProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAla 159
DB 512 TACCATCACAGCATCTGAGACTACTCGGCACGGCTAATAAGGCATCTAGTTTCAGCT 571
QY 160 IleGlnThrLeuPheThr---GlyProValLysAlaGluGluProThrProThrProLys 178
DB 572 ATCCAGAGCTTCTTCGCTGMAACCACTGTATAAGCTGAAGACACCAACCCCAAAA 631
QY 179 IleProLysLysArgGlyArgLysLysLysAlaAspProGluGluIleAsnSerSer 198
DB 632 GTTCTTAAAGAGAGAGGAGGAGGAAGAAGACGCTGATCTCTGAGGAATAAATTCATCT 691
QY 199 AlaProArg 201
DB 692 GCTCCACGA 700

RESULT 11
LOCUS BX835682/c 1041 bp mRNA linear EST 11-FEB-2004
DEFINITION BX835682 Arabidopsis thaliana Adult vegetative tissue Col-0
ACCESSION Arabidopsis thaliana cDNA clone GSLTUS48ZH11 3PRIM, mRNA sequence.
VERSION BX835682.1 GI:42529765

KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1041)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished (2004)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.

FEATURES
  source
    1..1041
      /organism="Arabidopsis thaliana"
      /mol_type="mRNA"
      /ecotype="Col-0"
      /db_xref="taxon:3702"
      /clone="GSLTUS48ZH11"
      /tissue_type="Adult vegetative tissue"
      /clone_lib="Arabidopsis thaliana Adult vegetative tissue
      Col-0"

ORIGIN
  Alignment Scores:
    Pred. No.: 1,21e-81 Length: 1041
    Score: 860.00 Matches: 177
    Percent Similarity: 79.10% Conservative: 16
    Best Local Similarity: 72.54% Mismatches: 31
    Query Match: 47.88% Indels: 20
    DB: 5 Gaps: 6

US-10-088-187A-11 (1-341) x BX835682 (1-1041)

QY 112 SerAlaHisAsnHisPheLysArg-----AlaArgLeu 122
DB 1009 ACTGCGCACACCACTTACTCGTCGTGACGCACTTTGACTTTTGGTTACCGCACAGTTA 950
QY 123 Phe-----GluAspLeuGluAspAlaGluValIlePheProSerSer 138
DB 949 CTGTACTACTAACGTACACGTTTGTCTGACAGCACTGACACACGCTT-----ACTAGG 896
QY 139 ValTyrProSerProLeuProGluSerThrValProAlaAsnLysGlyTyr---AlaSer 157
DB 895 GACATTTAAACACACGAAACTGAACTACTGAACCACTGAAAGGTTATGGCGGTTCT 836
QY 158 SerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrPro 177
DB 835 ACAGCCATCCAAAGCTTTTTCAAA---GAATCTAAAGCTGAAGAA-----ACGCC 788
QY 178 LysIleProLysLysArgGlyArgLysLysAsnAlaAspProGluGluIleAsnSer 197
DB 787 AAGGTACTTAAGAGAGAGAGAGAGAGAGAGAGATCTCTATCCGAGAGTAAGTAACTCT 728
QY 198 SerAlaProArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAla 217

```

```

Db      727  TCAACACCCGGTGGAGTCACTCAGAGAACCCGCTCAAGTCTTACGAGAGTCTTCTGCT 668
Qy      218  ArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPhe 237
Db      667  AGAAGAGAACTGTAACTCAGAGGAAAGAGAGAGCCGCTCAATGCAGCCAAACATTC 608
Qy      238  GluProThrAsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCys 257
Db      607  GAACCAACAACATCTTACTTTAGAGTTGTCTCGCACCACATCATATATACAGAGGTTGC 548
Qy      258  IleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIle 277
Db      547  ATCATGTACTTGCATCTGGTGTCTGAGAAATACCTAAGTGGGATATCTGGTTTCATC 488
Qy      278  LysValGlnLeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAla 297
Db      487  AGCTCCAGCTCGGTGAGAAACAATGGCCAGTGAGTGCTCTACNAGCAGGAGAGCT 428
Qy      298  LysPheSerGlnGlyTyrTrpTyrGluPheThrArgAspPheValLeuLysValThrAlaPheArgVal 317
Db      427  AAGTTTAGCAAGGATGGTATGAGTTTCACTCGAGAACATATATAGGCGAAGGAGATGA 368
Qy      318  CysValPheGluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgVal 337
Db      367  TGTGTGTTTGAGCTACTCAGAACTCGGATTTCTGTTCTCGAAGTCACCGCTTTCTGTGTC 308
Qy      338  AsnGluTyrVal 341
Db      307  AATGAGTATGTC 296

```

```

RESULT 12
LOCUS   CO108208
DEFINITION GR_Eb0039N11.r GR_Eb Gossypium raimondii cDNA clone GR_EB0039N11
3', mRNA sequence.
ACCESSION CO108208
VERSION   CO108208.1 GI:48806894
KEYWORDS EST.
SOURCE   Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 834)
Kim.H., Yu.Y., Kudrna,D., Hatfield,J., Strum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 0039 row: N column: 11.

```

```

FEATURES
source
1..834
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb0039N11"
/tissue_type="Floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/notes="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Clones
plated/picked by AGI. More glycerol clones held in -80."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      1.66e-76      Length:      834
Score:          811.00      Matches:    169
Percent Similarity: 73.9%      Conservative: 44
Best Local Similarity: 58.6%      Mismatches: 52
Query Match:    45.1%      Indels:    23
DB:             7          Gaps:      6

```

US-10-088-187A-11 (1-341) x CO108208 (1-834)

```

Qy      9  LeuIlePheSerSerThrIleGln-GluLysArgLeuArgValProAspLysPheValSe 28
Db      2  CTTATTCTCTTACCACCCCTCCAGCAAAATGAACTGAGGATCCCTGATAACTTCGTAA 61
Qy      28  rlyPheLysAspGluLeuSerValAlaValAlaLeuThrValProAspGlyHisValTr 48
Db      62  GAAATTCAGGAGTAACCTTTCTGTTGCTGCCCTCTCACTGTCTCTGACGGTCATGTTG 121
Qy      48  pArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGlyTrpGlnGluPh 68
Db      122  GCGTGTAGGAATAAAGAAAGTTGACACACAGGTTTGGTTTCAGGAAGGTTGCAGAGTT 181
Qy      68  eValAspArgTyrSerIleArgIleGlyTyrLeuLeuIlePheArgTyrGluGlyAsnSe 88
Db      182  TCTAGAGCGTTACTATATTCTGTTGGCTACGTATTGGTTTGTAGTACGAGGAATTC 241
Qy      88  rAlaPheSerValTyrIlePheAsnLeuSerHisSerGluIleAsnTyrHisSerThrCl 108
Db      242  TGCITTCAGTGTAGTATATTAAATTTGTACAACTCGGAATAATACTATCAGACTAATGC 301
Qy      108  yLeuMetAspSerAlaHisAsnHisPheLysArgAlaArgLeuPheGluAspLeuGluAs 128
Db      302  CCTCGTTGGTACTCAATACATCAGGAAACAAATATCCA---TTTGAACAACCTTGAAGA 358
Qy      128  pGluAspAla-----GluValIlePheProSerSerValTyrProSerPr 143
Db      359  TGATGAATGCATCTCTCCAGCACCTTCAGAAATTTGTTGGTGGTCTAAACCTTAACAAC 418
Qy      143  oLeu-----ProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaTrl 160
Db      419  CATAACTGGGGTGGTGATGCCAATCTTCAAACTTCAAGGGTGTGAATATATCAACCTAT 478
Qy      160  eGln-----ThrLeuPheThrGlyProValLysAlaGluGluProThrProth 176
Db      479  TCGAGTCAAAATGCATACTTCAGGTTTCAGTGCATGAACCCAGAA-----524
Qy      176  rProLysIleProLysLysArgGlyArgLysLysAsnAlaAspProGluGluLeAs 196
Db      525  -----CCAAAAAAGCGTGGGAGAAAGCGGAAG---TTTGATCCTTAACGTGCAGGA 571
Qy      196  nSerSerAlaProArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlase 216
Db      572  TTTCATCTCTCGAGCTGAAGATGATGTTGATTCGCTTTAGATGTTTAAAGTGCCTTC 631
Qy      216  rAlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysTh 236
Db      632  AGCCGAAAGAGAACCCGTGACAGCTGAAGAAGAGAGAGAGCAATCATATGAGCCAAAGC 691
Qy      236  rPheGluProThrAsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGl 256
Db      692  ATTTGAGCCAACTAACCCCTTTCTGCAGGGTCTGCTTTCGAGCCATCTTATCTGTACAGGG 751
Qy      256  yCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPh 276
Db      752  ATGTATTATGTACTTACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811
Qy      276  eIleLysValGlnLeuAlaGlu 283
Db      812  CATTAACCTTACGCTTCTCTGAT 833

```

RESULT 13
AV821829

LOCUS AV821829 669 bp mRNA linear EST 01-APR-2002
 DEFINITION AV821829 RAF14 Arabidopsis thaliana cDNA clone RAF104-16-P22 5', mRNA sequence.
 ACCESSION AV821829 GI:19863860
 VERSION AV821829.1
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 REFERENCE 1 (bases 1 to 669)
 AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.
 TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
 JOURNAL Unpublished (2002)
 COMMENT Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@tc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified phagescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES
 source
 1. 669
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF104-16-P22"
 /dev_stage="rosette plants"
 /lab_host="SOLR"
 /clone_lib="RAF14"
 /note="Site 1: SstI; Site 2: XhoI; subjected to cold-treated(1,2,5,10,24 hr)"

ORIGIN
 Alignment Scores:
 Pred. No.: 2 5e-74 Length: 669
 Score: 789.50 Matches: 171
 Percent Similarity: 79.75% Conservative: 18
 Best Local Similarity: 72.15% Mismatches: 22
 Query Match: 43.96% Indels: 26
 DB: 1 Gaps: 9

US-10-088-187a-11 (1-341) x AV821829 (1-669)

QY 51 GlyLeuArgLYeAlaaspAsnlyslftrPheGlnAspGlyTrpGlnclupheValasp 70
 DB 15 GGACTAAGCAAGCAACCAACAAATCTGTTTCAAGACGGTTGCGAGGAGTTGTCAAC 74
 QY 71 ArgTyrSerIleArgIleGlyTyrLeuLeuIlePheArgTyrGluGlyAsnSerAlaPhe 90
 DB 75 CGTTTCTCCATTCGGATTGGT-----TTCAAGATACAAA-----GTT 110
 QY 91 SerValTyrIlePheAsnLeu---Ser-HisSerGluIleAsnTyrHisSerThrGly-- 108
 DB 111 ACAGTCTACATTTTCAATTTATCTCCACACCTGAGATCAACACCATTCCTAGTAGTGA 170
 QY 109 -----LeuMetAspSerAlaHisAsnHisPhe---LysArgAlaArgLeuPheGl 124
 DB 171 AGCTCTTATGCAATGGATTCCGACAGAAATCAGTTCACAAACGGTTCGATGTTTGA 230
 QY 124 uAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyrProSerProle 144
 DB 231 AGATCTCTGAACCAAGATGCTAAGGTCAATTATTCATCGAAC----- 273

QY 144 uProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAla---IleGlnThrLe 163
 DB 274 -CCTGATCTACTGAACCAAGTGAATAAGGTTATGGCGGTTCTACAGACATCCAAAGCTT 332
 QY 163 uPheThrGlyProValLysAlaGluGluProThrProThrProLysIleProLysLysAr 183
 DB 333 TTTCAAA---GAATCTAAAGCTGAAGAA-----ACGCCCAAGGTACTTTAAGAAGAG 380
 QY 183 qGlyValGlyLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAspAs 203
 DB 381 AGHAGAGAGAGAGAGAGAGTCTTAATCCCGAGAGAGTAAATCTTCAACTCCCGTGGAGA 440
 QY 203 pAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrValTh 223
 DB 441 TGACTCAGAGAACCGCTCAAGTTCTACGAGAGTGTCTTCTGTAGAAAGAGAACTGTAAC 500
 QY 223 rAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThrAsnProPh 243
 DB 501 TGCAGAGAGAGAGAGAGAGCGCTCAATGCGAGCAAAACATTCGAAACCAATCCTTA 560
 QY 243 ePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeuProSe 263
 DB 561 CTTTAGAGATTGTTCTGCGACCATCATATCATATACAGAGGTTCATCATGTACTTGCATC 620
 QY 263 rGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysVal 279
 DB 621 TGGGTTTGTGAGAAATACCTAAGTGGATATCTATCATGAGGTTCATCATCAAGCTC 669

RESULT 14
 CD823066 645 bp mRNA linear EST 10-JUL-2003
 LOCUS BN25.047120F020109 BN25 Brassica napus cDNA clone BN25047120, mRNA sequence.
 DEFINITION CD823066
 ACCESSION CD823066
 VERSION CD823066.1 GI:32505006
 KEYWORDS EST.
 SOURCE Brassica napus (rape)
 ORGANISM Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 645)
 AUTHORS Genoplatte.
 TITLE Genoplatte, a major partnership french program in plant genomics
 JOURNAL Unpublished (2003)
 COMMENT Contact: Genoplatte
 Genoplatte
 93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french plant genomics programme 'Genoplatte' (<http://www.genoplatte.com>) and <http://genoplatte-info.infobiogen.fr>.

FEATURES
 source
 1. 645
 /organism="Brassica napus"
 /mol_type="mRNA"
 /cultivar="Jet Neuf"
 /db_xref="taxon:3708"
 /clone="BN25047120"
 /tissue_type="seed"
 /clone_lib="BN25"

ORIGIN
 Alignment Scores:
 Pred. No.: 5 66e-74 Length: 645
 Score: 786.00 Matches: 149
 Percent Similarity: 98.70% Conservative: 3
 Best Local Similarity: 96.75% Mismatches: 2
 Query Match: 43.76% Indels: 0
 DB: 6 Gaps: 0

US-10-088-187a-11 (1-341) x CD823066 (1-645)

Search completed: December 30, 2004, 09:49:31
Job time : 6758 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2004, 13:08:18 ; Search time 4772 Seconds
(without alignments)
10177.393 Million cell updates/sec

Title: US-10-088-187A-10_COPY_269_1295
Perfect score: 1027
Sequence: 1 atgccacgcccctttctcca.....agtcacgagtagctgtctgaa 1027

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1027	100.0	1495	6 AX101036	AX101036 Sequence
2	1027	100.0	1538	8 AF289052	AF289052 Arabidops
3	1026	99.9	1026	6 CO805406	CO805406 Sequence
4	1025.4	99.8	1495	6 AX101038	AX101038 Sequence
5	1015	98.8	1494	6 AX101039	AX101039 Sequence
6	673	65.5	990	8 AY356368	AY356368 Brassica
7	671.4	65.4	990	8 AY517929	AY517929 Brassica
8	653.4	63.6	5000	8 AF289051	AF289051 Arabidops
9	653.4	63.6	79186	8 AP000735	AP000735 Arabidops
10	653	63.6	5000	6 AX101027	AX101027 Sequence
11	442.6	43.1	1146	6 AY042868	AY042868 Arabidops
12	397.6	38.7	681	6 AX505578	AX505578 Sequence
13	270.4	26.3	125021	8 AC007504	AC007504 Arabidops
14	140.4	13.7	78589	8 AP006145	AP006145 Lotus cor
15	132.6	12.9	243777	2 AC137825	AC137825 Medicago
16	101	9.8	2000	6 AX508278	AX508278 Sequence
17	89	8.7	48368	2 AC140105	AC140105 Medicago
18	83.4	8.1	114098	8 AC144459	AC144459 Medicago
19	81.2	7.9	108028	8 AC135233	AC135233 Medicago

20	80.8	7.9	48368	2	AC140105	AC140105 Medicago
21	72.8	7.1	129204	8	AC146747	AC146747 Medicago
22	65	6.3	92281	8	AP004896	AP004896 Lotus cor
23	62.2	6.1	111252	8	AC148404	AC148404 Medicago
24	61.6	6.0	573	8	BT010604	BT010604 Arabidops
25	61.2	6.0	661	8	BT008541	BT008541 Arabidops
26	61.2	6.0	809	8	BT006139	BT006139 Arabidops
27	56.4	5.5	122693	8	AC145449	AC145449 Medicago
28	56.2	5.5	204120	8	AY661659	AY661659 Sorghum b
29	51.8	5.0	129647	8	AC097280	AC097280 Oryza sat
30	51.8	5.0	140383	8	CNS08CD5	AL954829 Oryza sat
31	50.8	4.9	90596	8	TI5B16	AF104919 Arabidops
32	50.8	4.9	198697	8	ATCHRIV4	AL161492 Arabidops
33	50	4.9	125021	8	AC007504	AP007504 Arabidops
34	48	4.7	79186	8	AP000735	AP000735 Arabidops
35	46.4	4.5	118966	2	AC147012	AC147012 Medicago
36	45.2	4.4	67	8	ATH526944	AJ56944 Arabidops
37	43.8	4.3	1463	8	ATH441294	AJ441294 Arabidops
38	43.8	4.3	90902	8	ATF28M20	AL031004 Arabidops
39	43.8	4.3	196242	8	ATCHRIV75	AL161579 Arabidops
40	43.6	4.2	226320	2	AC118481	AL118481 Rattus no
41	43.2	4.2	7218	6	I66494	I66494 Sequence 14
42	43.2	4.2	206893	2	AC115569	AC115569 Rattus no
43	43.2	4.2	270068	2	AC113213	AC113213 Rattus no
44	43	4.2	832	6	AR415176	AR415176 Sequence
45	43	4.2	832	6	AX972010	AX972010 Sequence

ALIGNMENTS

RESULT 1
AX101036
LOCUS AX101036 1495 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 10 from Patent WO0121822.
ACCESSION AX101036
VERSION AX101036.1 GI:13619892

KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1
AUTHORS Dean, C. and Levy, Y. Y.
TITLES Methods and means for modification of plant flowering characteristics
JOURNAL Patent: WO 0121822-A 10 29-MAR-2001;
Plant Bioscience Limited (GB)

FEATURES
Location/Qualifiers
1..1495
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
269..1294
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC36501.1"
/db_xref="GI:13619893"

CDS
269..1294
/translation="MNPFFHKLIFSSITQEKRLRVPDKFVSKFDELSVAVALTVPD
GHVWVGLRKADNKIWFOGWOEFVDYRSIRIGYLLIFRYEGNSAFSVYIFNLSHSEI
NYHSLGMDSAHNHFKEARLDELDEDAEVI FPPSSVPSPLPESTVPANKGVASSAI
QTLTGVPKAEPTTPIPKKGRKKKNADPEEINSSAPRDDDRNSKFFVESASAR
KKTVAERERAINAKTFEPTNFFRVLRPSYLYRGCMITLPLPSGFAKYLUSIGSGF
IKVQAELQKQWPCVRCCLYKAGRAKFSQGWYFTLENNLGEDVCVFEELLTRDFVLKTA
FRVNEYV"

ORIGIN
Query Match 100.0%; Score 1027; DB 6; Length 1495;
Best Local Similarity 100.0%; Pred. No. 2.3e-287;
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCCACGCCCTTCTTCTCCATAAGTTGATTTTCTCATCCACTATCCAGAAAAACGCTG 60

Db 269 ATGCCACGCCCTTTCTCCATAGTTGATTTTCTCATCCATATCCAGAAAAAGCTGCTG 328
Qy 61 AGGGTCCAGATAAGTTGTGAGTAAATTCAGAGATGAGCTTTCGGTGTGTTGCATC 120
Db 329 AGGGTCCAGATAAGTTGTGAGTAAATTCAGAGATGAGCTTTCGGTGTGTTGCATC 388
Qy 121 ACAGTACCTGATGGTTCATGTTTGGCGGTGAGACTAAGAAAGCTGACAAACAAAATTTGG 180
Db 389 ACAGTACCTGATGGTTCATGTTTGGCGGTGAGACTAAGAAAGCTGACAAACAAAATTTGG 448
Qy 181 TTTCAGATGTTGGCAAGATTTGTTGACCGTTACTCAATTCGATCGTATCTTTTG 240
Db 449 TTTCAGATGTTGGCAAGATTTGTTGACCGTTACTCAATTCGATCGTATCTTTTG 508
Qy 241 ATTTTATGATGAGGAACCTCGCTTCAGCTACATTTTCAATTTATCCACTCT 300
Db 509 ATTTTATGATGAGGAACCTCGCTTCAGCTACATTTTCAATTTATCCACTCT 568
Qy 301 GAGATCAATTACATTCACCGGTCTCATGATTCGGCTCAACACACTTCAACGCGCC 360
Db 569 GAGATCAATTACATTCACCGGTCTCATGATTCGGCTCAACACACTTCAACGCGCC 628
Qy 361 CGTTGTTTGAAGACTTGAAGATGAAGATCCGAGGTCACTTTCTTCTTCTGTGATC 420
Db 629 CGTTGTTTGAAGACTTGAAGATGAAGATCCGAGGTCACTTTCTTCTTCTGTGATC 688
Qy 421 CCATCACCCTTCCTGAGTCTACAGTCCAGCCCAAGAGGTGATGATTCAGCCATC 480
Db 689 CCATCACCCTTCCTGAGTCTACAGTCCAGCCCAAGAGGTGATGATTCAGCCATC 748
Qy 481 CAAACCTTGTTCACCTGGACCACTTAAAGCTGAAGAGCCAAAGCCCAACCCCAAAATACCT 540
Db 749 CAAACCTTGTTCACCTGGACCACTTAAAGCTGAAGAGCCAAAGCCCAACCCCAAAATACCT 808
Qy 541 AAAAAGAGAGGAGGAAGAAAATGCTGATCTCTGAGGAAATAAATCACTCAGCTCCG 600
Db 809 AAAAAGAGAGGAGGAAGAAAATGCTGATCTCTGAGGAAATAAATCACTCAGCTCCG 868
Qy 601 CGAGATGATGATCCAGAGAACCGTTCAAGTCTACGAGAGTCTTCTCGGAGAAAGAGA 660
Db 869 CGAGATGATGATCCAGAGAACCGTTCAAGTCTACGAGAGTCTTCTCGGAGAAAGAGA 928
Qy 661 ACCGTGACTGCAGAGAAAGAGAGAGAGCCATCAATGAGCCAAAGAGCTTCAAGCAACA 720
Db 929 ACCGTGACTGCAGAGAAAGAGAGAGAGCCATCAATGAGCCAAAGAGCTTCAAGCAACA 988
Qy 721 AACCTTTTCTCAGAGTGGTCTCGGACCATCTATCTATACAGAGGTTGCATCATGAT 780
Db 989 AACCTTTTCTCAGAGTGGTCTCGGACCATCTATCTATACAGAGGTTGCATCATGAT 1048
Qy 781 CTTCTCTTGGGTTTGTGAGAGTACCTTAAGTGGATCTCGGTTTCATCAAGTCCAG 840
Db 1049 CTTCTCTTGGGTTTGTGAGAGTACCTTAAGTGGATCTCGGTTTCATCAAGTCCAG 1108
Qy 841 CTTGCGGAGAAACATGGCTGTTCGATCTCTCAAAAGCCGAGAGCCAAATTCAGT 900
Db 1109 CTTGCGGAGAAACATGGCTGTTCGATCTCTCAAAAGCCGAGAGCCAAATTCAGT 1168
Qy 901 CAAGGATGATGAGAAATTCCTCTAGAGAAACAATTAGGAGAAAGGAGAGCTGTGTGTTT 960
Db 1169 CAAGGATGATGAGAAATTCCTCTAGAGAAACAATTAGGAGAAAGGAGAGCTGTGTGTTT 1228
Qy 961 GAGTGTCTCAGAACCGAGATTTGTTTGAAGTGAAGCCCTTCGAGTCAACGAGTAC 1020
Db 1229 GAGTGTCTCAGAACCGAGATTTGTTTGAAGTGAAGCCCTTCGAGTCAACGAGTAC 1288
Qy 1021 GTCTGAA 1027
Db 1289 GTCTGAA 1295

AF289052 1538 bp mRNA linear PLN 05-NOV-2003
LOCUS Arabidopsis thaliana reduced vernalization response 1 (VRN1) mRNA,
DEFINITION VRN1-Ler allele, complete cds.
ACCESSION AF289052
VERSION AF289052.1 GI:21734795
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
REFERENCE 1 (bases 1 to 1538)
AUTHORS Levy,Y., Mesnage,S., Mylne,J.S., Gendall,A.R. and Dean,C.
TITLE Multiple roles of Arabidopsis VRN1 in vernalization and flowering
time control
JOURNAL Science 297 (5579), 243-246 (2002)
MEDLINE 22111275
PUBMED 12114624
REFERENCE 2 (bases 1 to 1538)
AUTHORS Levy,Y., Gendall,A.R. and Dean,C.
TITLE VRN1, a gene required for response to vernalization
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1538)
AUTHORS Levy,Y. and Dean,C.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2000) Molecular Genetics, John Innes Centre,
Colney Lane, Norwich, Norfolk NR4 7UH, UK
FEATURES
source
1..1538
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/map="between mi339 and m560B2"
/ecotype="Landsberg erecta"
1..1538
/gene="VRN1"
/allele="Ler"
312..1337
/gene="VRN1"
/allele="Ler"
/note="contains two VPI/ABI3-like B3 domains"
/product="reduced vernalization response 1"
/protein_id="AA076973.1"
/db_xref="GI:21734795"
/translation="MPRPFFHKLIFSSTIQKRLRVPDKFVKDELVSVALVTVPD
GHVVRGLKADNKIWFQDQWQEFVDYRYSIRIGYLLIFRYEGNSAFSVYIENLHSEI
NYHSTGLMDSAHNHFKRALPDELDEDAEVIFFPSVYFPLPSTVPANKFYASSAI
QTLFTGPVKAEPPTPKIPKGRKKKNADPEEINSAPRDDDPENRKFVESASAR
KRTVABEREBAINAAKTFEPTNPFVRLPSYLRGCIIVLPSPGAEKYLSGTSGF
IKVQLAEKQWPRCLYKAGRAKFSQGWYEFLENNLGEQVCFVPELLTRDFVLKVTA
FRVNEYV"
ORIGIN
Query Match 100.0%; Score 1027; DB 8; Length 1538;
Best Local Similarity 100.0%; Pred. No. 2.3e-287;
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCCACGCCCTTTCTCCATAAGTTGATTTTCTCATCCACTATCCAGAAAAAGCTGCTG 60
Db 312 ATGCCACGCCCTTTCTCCATAAGTTGATTTTCTCATCCACTATCCAGAAAAAGCTGCTG 371
Qy 61 AGGGTCCAGATAAGTTGTGAGTAAATTCAGAGATGAGCTTTCGGTGTGTTGCATC 120
Db 372 AGGGTCCAGATAAGTTGTGAGTAAATTCAGAGATGAGCTTTCGGTGTGTTGCATC 431
Qy 121 ACAGTACCTGATGGTTCATGTTTGGCGGTGAGACTAAGAAAGCTGACAAACAAAATTTGG 180
Db 432 ACAGTACCTGATGGTTCATGTTTGGCGGTGAGACTAAGAAAGCTGACAAACAAAATTTGG 491
Qy 181 TTTCAGATGTTGGCAAGATTTGTTGACCGTTACTCAATTCGATCGTATCTTTTG 240

492 TTTCAAGATGGTGGCAAGATTTGTTGACCGTTACTCCATTCGCAATGGATTATCTTTG 551
241 ATTTTATAGATATGAAGAAACTCTGCTTTCAGCGTCTACATTTTCAATTTATCCCACTCT 300
552 ATTTTATAGATATGAAGAAACTCTGCTTTCAGCGTCTACATTTTCAATTTATCCCACTCT 611
301 GAGATCAATTAACATTCACCGGCTCTCATGGATTCGGTCTCAACACCACTTCAAAACGCGC 360
612 GAGATCAATTAACATTCACCGGCTCTCATGGATTCGGTCTCAACACCACTTCAAAACGCGC 671
361 CGTTTGTGTTGAGACCTTGAAGATGAAGATGCGGAGTTCATCTTTCTTTCTGTTGATC 420
672 CGTTTGTGTTGAGACCTTGAAGATGAAGATGCGGAGTTCATCTTTCTTTCTGTTGATC 731
421 CCATCAACCACTTCTGAGTCTACAGTACCAAGCAACAAAGGGTATGCTAGTTTCAGCCATC 480
732 CCATCAACCACTTCTGAGTCTACAGTACCAAGCAACAAAGGGTATGCTAGTTTCAGCCATC 791
481 CAAACCTTGTTCATCTGGAACAGTTAAAGCTGAAGAGCCAAACGCAACCCCAAAATACCT 540
792 CAAACCTTGTTCATCTGGAACAGTTAAAGCTGAAGAGCCAAACGCAACCCCAAAATACCT 851
541 AAAAGAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
852 AAAAGAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 911
601 CGAGATGATGATCCAGAGAACCGTTCAAAAGTTCTACGAGAGTCTTCGCGAGAGAGAGA 660
912 CGAGATGATGATCCAGAGAACCGTTCAAAAGTTCTACGAGAGTCTTCGCGAGAGAGAGA 971
661 ACCGTGACTGCAG 720
972 ACCGTGACTGCAG 1031
721 AACCCCTTCTCAGAGTGGTCTCGAGCAATCTATCTATACAGAGGTTGATCATGTAT 780
1032 AACCCCTTCTCAGAGTGGTCTCGAGCAATCTATCTATACAGAGGTTGATCATGTAT 1091
781 CTTCTCTTCTGGTGTCTCAGAGTACCTAAGTGGATCTCCGGTTCATCAAAAGTCCAG 840
1092 CTTCTCTTCTGGTGTCTCAGAGTACCTAAGTGGATCTCCGGTTCATCAAAAGTCCAG 1151
841 CTTGCGGAGAAACAAATGGCTCTTCGATCTCTACAAAGCGGAGAGAGAGAGAGAGAGAG 900
1152 CTTGCGGAGAAACAAATGGCTCTTCGATCTCTACAAAGCGGAGAGAGAGAGAGAGAGAG 1211
901 CAAAGATGTACGAATTCATCTAGAGAAACATTTAGAGAGAGAGAGAGAGAGAGAGAGAG 960
1212 CAAAGATGTACGAATTCATCTAGAGAAACATTTAGAGAGAGAGAGAGAGAGAGAGAGAG 1271
961 GAGCTGCTCAGAACCAAGATTTCTGTTTGAAGTGACAGCTTTGAGTCAACAGATAC 1020
1272 GAGCTGCTCAGAACCAAGATTTCTGTTTGAAGTGACAGCTTTGAGTCAACAGATAC 1331
1021 GTCTGAA 1027
1332 GTCTGAA 1338

RESULT 3
LOCUS CQ805406 1026 bp DNA linear PAT 10-MAY-2004
DEFINITION Sequence 1817 from Patent WO2004035798.
ACCESSION CQ805406
VERSION CQ805406.1 GI:47111337
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1

AUTHORS Inze,D., de Veylder,L. and Vlieghe,K.
TITLE Identification of novel e2f target genes and use thereof
JOURNAL Patent: WO 2004035798-A 1817 29-APR-2004;
CropDesign N.V. (BE)
FEATURES
source Location/Qualifiers
1..1026
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 99.9%; Score 1026; DB 6; Length 1026;
Best Local Similarity 100.0%; Pred. No. 4.3e-287;
Matches 1026; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCCACGCCCTTCTTCATTAAGTTGATTTTCTCATCCACTATCCACGAAAAACGCTCG 60
DB 1 ATGCCACGCCCTTCTTCATTAAGTTGATTTTCTCATCCACTATCCACGAAAAACGCTCG 60
QY 61 AGGGTCCAGATAAGTTTGTGAGTAAATTTCAAGATGAGCTTTCGGTTCGCTGTCACATC 120
DB 61 AGGGTCCAGATAAGTTTGTGAGTAAATTTCAAGATGAGCTTTCGGTTCGCTGTCACATC 120
QY 121 ACAGTACCTGATGCTGTCATGTTTGGCGTGTAGACTAAGGAAAGCTGACAAATAATTTGG 180
DB 121 ACAGTACCTGATGCTGTCATGTTTGGCGTGTAGACTAAGGAAAGCTGACAAATAATTTGG 180
QY 181 TTTCAAGATGTTGGCAAGAGTTTCTTGACCGTTACTCCATTCGCAATGGTTATCTTTTG 240
DB 181 TTTCAAGATGTTGGCAAGAGTTTCTTGACCGTTACTCCATTCGCAATGGTTATCTTTTG 240
QY 241 ATTTTATAGATGAAGGAAACTCTGCCCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 300
DB 241 ATTTTATAGATGAAGGAAACTCTGCCCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 300
QY 301 GAGATCAATTAACATTCACCGGCTCTCATGGATTCGGTCTCAACACCTTCAAAACGCGCC 360
DB 301 GAGATCAATTAACATTCACCGGCTCTCATGGATTCGGTCTCAACACCTTCAAAACGCGCC 360
QY 361 CGTTTGTGTTGAGACCTTGAAGATGAAGTCCGAGGTCATCTTCTCTTCTGTTGATC 420
DB 361 CGTTTGTGTTGAGACCTTGAAGATGAAGTCCGAGGTCATCTTCTCTTCTGTTGATC 420
QY 421 CCATCAACCACTTCTGAGTCTACAGTACCGAGCAACAAAGGGTATGCTAGTTTCAGCCATC 480
DB 421 CCATCAACCACTTCTGAGTCTACAGTACCGAGCAACAAAGGGTATGCTAGTTTCAGCCATC 480
QY 481 CAAACCTTGTTCATGGAACCGATTAAAGCTGAAGAGCCAAACGCAACCCCAAAATACCT 540
DB 481 CAAACCTTGTTCATGGAACCGATTAAAGCTGAAGAGCCAAACGCAACCCCAAAATACCT 540
QY 541 AAAAGAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 541 AAAAGAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 601 CGAGATGATGATCCAGAGAAACCGTTTCAAAAGTTCTACGAGAGTCTTTCGCGAGAAAGAGA 660
DB 601 CGAGATGATGATCCAGAGAAACCGTTTCAAAAGTTCTACGAGAGTCTTTCGCGAGAAAGAGA 660
QY 661 ACCGTGACTGCAGAGAAAG 720
DB 661 ACCGTGACTGCAGAGAAAG 720
QY 721 AACCCCTTCTCAGAGTGGTTCGCGACCATCTATCTATACAGAGGTTGATCATGTAT 780
DB 721 AACCCCTTCTCAGAGTGGTTCGCGACCATCTATCTATACAGAGGTTGATCATGTAT 780
QY 781 CTTCTCTTCTGGTGTGCTGAGAGTACCTAAGTGGATCTCCGGTTCATCAAAAGTCCAG 840
DB 781 CTTCTCTTCTGGTGTGCTGAGAGTACCTAAGTGGATCTCCGGTTCATCAAAAGTCCAG 840
QY 841 CTTGCGGAGAAACAAATGGCTCTTCGATGTCTCTCAAAAGCGGAGAGAGAGAGAGAGAG 900
DB 841 CTTGCGGAGAAACAAATGGCTCTTCGATGTCTCTCAAAAGCGGAGAGAGAGAGAGAGAG 900

```
Db      841  CTTGCGGAGAAACAATGGCGTGTTCGATGCTCTCAAAAGCGGGAGAGCAAAATTCAGT 900
Qy      901  CRAAGGTGTACGAATTCCTCTAGAGAACACTTAGGAGAGGAGACCTCTGTGTGTTT 960
Db      901  CRAAGGTGTACGAATTCCTCTAGAGAACACTTAGGAGAGGAGACCTCTGTGTGTTT 960
Qy      961  GAGCTGCTCAGAACACAGAGATTTCTGTTTGAAGTGCACAGCCCTTTCCAGTCAACAGGTAC 1020
Db      961  GAGCTGCTCAGAACACAGAGATTTCTGTTTGAAGTGCACAGCCCTTTCCAGTCAACAGGTAC 1020
Qy      1021  GTCTGA 1026
Db      1021  GTCTGA 1026

RESULT 4
AX101038
LOCUS      AX101038      1495 bp      DNA      linear      PAT 10-APR-2001
DEFINITION Sequence 12 from Patent WO0121822.
ACCESSION AX101038
VERSION    AX101038.1 GI:13619894
KEYWORDS   .
SOURCE      synthetic construct
            synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS     Dean,C. and Levy,Y.Y.
TITLE        Methods and means for modification of plant flowering
            Characteristics
JOURNAL      Patent: WO 0121822-A 12 29-MAR-2001;
            Plant Bioscience Limited (GB)
FEATURES     Location/Qualifiers
            1..1495
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="vrnl-1 mutation"

ORIGIN
Query Match      99.8%; Score 1025.4; DB 6; Length 1495;
Best Local Similarity 99.9%; Pred. No. 6.7e-287;
Matches 1026; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  ATGCCACGCCCTTCTTCCATAAGTTGATTTCTCATCCACTATCCAGAAAAAGCTCG 60
Db      269  ATGCCACGCCCTTCTTCCATAAGTTGATTTCTCATCCACTATCCAGAAAAAGCTCG 328
Qy      61  AGGTCCTCCAGATAAGTTGTGAGTAAATTCAGGATGAGCTTTCGGTTGCTGTGCACATC 120
Db      329  AGGTCCTCCAGATAAGTTGTGAGTAAATTCAGGATGAGCTTTCGGTTGCTGTGCACATC 388
Qy      121  ACAGTACCTGATGGTCATCTTTGGCGTGTAGGACTAAGGAAAGCTGACAAACAAATTTGG 180
Db      389  ACAGTACCTGATGGTCATCTTTGGCGTGTAGGACTAAGGAAAGCTGACAAACAAATTTGG 448
Qy      181  TTTCAAGATGGTTGGCAAGATTTGTGTACCGTTACTCCATTGCGATGGTTATCTTTTG 240
Db      449  TTTCAAGATGGTTGGCAAGATTTGTGTACCGTTACTCCATTGCGATGGTTATCTTTTG 508
Qy      241  ATTTTATAGATAGAGAAACTCTGCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 300
Db      509  ATTTTATAGATAGAGAAACTCTGCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 568
Qy      301  GAGATCAATTACCATTTCCACCGCTCTCATGGATTCGGCTCAACACACTTCAACCGGCC 360
Db      569  GAGATCAATTACCATTTCCACCGCTCTCATGGATTCGGCTCAACACACTTCAACCGGCC 628
Qy      361  CGTTTCTTTGAAGACCTTGAAGATGAAGTGCAGGTCATCTTTCTCTTCTCTGTGTAC 420
Db      629  CGTTTCTTTGAAGACCTTGAAGATGAAGTGCAGGTCATCTTTCTCTTCTCTGTGTAC 688
Qy      421  CCAATCACCACTTCTGAGTCTCAGTACCAAGCAAAAGGATGCTAGTTTCAGGCATC 480
```

```
Db      689  CCATCACCACCTTCTCTGAGTCTACAGTACGACCAACAAAAGGATGTAGTTTCAGCCATC 748
Qy      481  CAAACCTTTGTTCACTGGACCAAGTTAAAGCTGAAGGCCAACGCCAACCCCAAAATACCT 540
Db      749  CAAACCTTTGTTCACTGGACCAAGTTAAAGCTGAAGGCCAACGCCAACCCCAAAATACCT 808
Qy      541  AAAAAGAGAGGAGGAGGAAGAAAATGCTGATCCTGAGGAAATATAAATCTCATCAGCTCG 600
Db      809  AAAAAGAGAGGAGGAGGAAGAAAATGCTGATCCTGAGGAAATATAAATCTCATCAGCTCG 868
Qy      601  CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTGTCTTCGAGAAAGAGA 660
Db      869  CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTGTCTTCGAGAAAGAGA 928
Qy      661  ACCGTGACTGCAGAGAAAGAGAGAGGCCATCAATGACGACCAAAACCTTCGAACCAACA 720
Db      929  ACCGTGACTGCAGAGAAAGAGAGAGGCCATCAATGACGACCAAAACCTTCGAACCAACA 988
Qy      721  AACCCCTTTCTTCAGAGTGGTTCTGCGACCATCTCTATCTATACAGAGGTTGCATCATGTAT 780
Db      989  AACCCCTTTCTTCAGAGTGGTTCTGCGACCATCTCTATCTATACAGAGGTTGCATCATGTAT 1048
Qy      781  CTTCTCTCTGGGTTTGTGAGAAATCACTAAGTGGATCTCCGGTTTCATCAAAAGTCCAG 840
Db      1049  CTTCTCTCTGGGTTTGTGAGAAATCACTAAGTGGATCTCCGGTTTCATCAAAAGTCCAG 1108
Qy      841  CTTGCGGAGAAACAATGGCGTGTTCGATGCTCTCAAAAGCGGGAGAGCCAAATTCAGT 900
Db      1109  CTTGCGGAGAAACAATGGCGTGTTCGATGCTCTCAAAAGCGGGAGAGCCAAATTCAGT 1168
Qy      901  CAAGGATGATGAGAAATTCAGTCTAGAGAAACAATTCAGGAGAGGAGAGCTGTGTGTTT 960
Db      1169  CAAGGATGATGAGAAATTCAGTCTAGAGAAACAATTCAGGAGAGGAGAGCTGTGTGTTT 1228
Qy      961  GAGCTGCTCAGAACACAGAGATTTGTTTGAAGTGCACAGCTTTCGAGTCAACAGGTAC 1020
Db      1229  GAGCTGCTCAGAACACAGAGATTTGTTTGAAGTGCACAGCTTTCGAGTCAACAGGTAC 1288
Qy      1021  GTCTGA 1027
Db      1289  GTCTGA 1295

RESULT 5
AX101039
LOCUS      AX101039      1494 bp      DNA      linear      PAT 10-APR-2001
DEFINITION Sequence 13 from Patent WO0121822.
ACCESSION AX101039
VERSION    AX101039.1 GI:13619895
KEYWORDS   .
SOURCE      synthetic construct
            synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS     Dean,C. and Levy,Y.Y.
TITLE        Methods and means for modification of plant flowering
            Characteristics
JOURNAL      Patent: WO 0121822-A 13 29-MAR-2001;
            Plant Bioscience Limited (GB)
FEATURES     Location/Qualifiers
            1..1494
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="vrnl-2 mutation"

ORIGIN
Query Match      98.8%; Score 1015; DB 6; Length 1494;
Best Local Similarity 99.9%; Pred. No. 7.1e-284;
Matches 1026; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy      1  ATGCCACGCCCTTCTTCCATAAGTTGATTTCTCATCCACTATCCAGAAAAAGCTCG 60
```

Db 269 ATGCCAGGCCCTTTCTTCCATAAGTTGATTTTCTCATCCATATCCAGAAAAACGCTCTG 328
QY 61 AGGGTCCAGATAAGTTTGTGAGTAAATTCAGAGATGAGCTTTCCGTTGCTGTGCACTC 120
Db 329 AGGGTCCAGATAAGTTTGTGAGTAAATTCAGAGATGAGCTTTCCGTTGCTGTGCACTC 388
QY 121 ACAGTACTGATGGTCATGTTTGGCGTGTAGGACTAAGGAAAGCTGACACAAAAATTTGG 180
Db 389 ACAGTACTGATGGTCATGTTTGGCGTGTAGGACTAAGGAAAGCTGACACAAAAATTTGG 448
QY 181 TTTCAAGATGTTGGCAAGATTTGTGACCGTTACTCCATTCGCAATGCTGTTATCTTTTGG 240
Db 449 TTTCAAGATGTTGGCAAGATTTGTGACCGTTACTCCATTCGCAATGCTGTTATCTTTTGG 508
QY 241 ATTTTATAGATATCAAGAAACTCTGCGCTTCAGCGTCTACATTTTCAATTTATCCACTCT 300
Db 509 ATTTTATAGATATCAAGAAACTCTGCGCTTCAGCGTCTACATTTTCAATTTATCCACTCT 568
QY 301 GAGATCAATTAACATTCACCGTCTCATGGATTCGCTCACAACCACTTCAAAACGCGC 360
Db 569 GAGATCAATTAACATTCACCGTCTCATGGATTCGCTCACAACCACTTCAAAACGCGC 628
QY 361 CGTTTCTTTGAACACCTTGAAGATGAAGATGCGCGAGTCACTTTCTTCTTCTGTTGATC 420
Db 629 CGTTTCTTTGAACACCTTGAAGATGAAGATGCGCGAGTCACTTTCTTCTTCTGTTGATC 688
QY 421 CCATCACCACTTCTCTGAGTCTACAGTACCAAGCAACAAAGGGTATGCTAGTTTCAGCGCATC 480
Db 689 CCATCACCACTTCTCTGAGTCTACAGTACCAAGCAACAAAGGGTATGCTAGTTTCAGCGCATC 748
QY 481 CAAACCTTGTCTCTGAGCAGATTAAGTGAAGAGCCAAACGCAACCCCAAAAAATACCT 540
Db 749 CAAACCTTGTCTCTGAGCAGATTAAGTGAAGAGCCAAACGCAACCCCAAAAAATACCT 808
QY 541 AAAAGAGAGGGAGGAGAAATGCTGATCTCTGAGCAATTAACCTCATCAGCTCCG 600
Db 809 AAAAGAGAGGGAGGAGAAAG-AAAATGCTGATCTCTGAGCAATTAACCTCATCAGCTCCG 867
QY 601 CGAGATGATGATCCAGAGAACCGTTCAAAGTCTACGAGATGCTTCTCGGAGAAAGAGA 660
Db 868 CGAGATGATGATCCAGAGAACCGTTCAAAGTCTACGAGATGCTTCTCGGAGAAAGAGA 927
QY 661 ACCGTGACTGAGAGAAAGAGAGAGAGCCATCAATGCGAGCCAAACGTTGCAACCAACA 720
Db 928 ACCGTGACTGAGAGAAAGAGAGAGAGCCATCAATGCGAGCCAAACGTTGCAACCAACA 987
QY 721 AACCCCTTCTTCAAGTGGTCTTCCGACCATCTATCTATACAGAGGTTGCATCATGTAT 780
Db 988 AACCCCTTCTTCAAGTGGTCTTCCGACCATCTATCTATACAGAGGTTGCATCATGTAT 1047
QY 781 CTTCTCTTCTGGGTTTCTGAGAGATACCTAAGTGGATCTCCGGGTTCTATCAAGTCCAG 840
Db 1048 CTTCTCTTCTGGGTTTCTGAGAGATACCTAAGTGGATCTCCGGGTTCTATCAAGTCCAG 1107
QY 841 CTTGCGGAGAAACAAATGGCTGTTCGATGTCTCTACAAAGCCGGAGAGCCAAATTCAGT 900
Db 1108 CTTGCGGAGAAACAAATGGCTGTTCGATGTCTCTACAAAGCCGGAGAGCCAAATTCAGT 1167
QY 901 CAAGATGTTAGCAATTCATCTAGAGAACCACTTAGGAGAGGAGAGCGTCTGTGTGTTT 960
Db 1168 CAAGATGTTAGCAATTCATCTAGAGAACCACTTAGGAGAGGAGAGCGTCTGTGTGTTT 1227
QY 961 GAGCTGCTCAGAACACAGAGATTTCTGTTTGAAGTGCAGAGCCCTTCCAGTCAACGAGTAC 1020
Db 1228 GAGCTGCTCAGAACACAGAGATTTCTGTTTGAAGTGCAGAGCCCTTCCAGTCAACGAGTAC 1287
QY 1021 GTCTGAA 1027
Db 1288 GTCTGAA 1294

LOCUS AY356368 990 bp mRNA linear PLN 25-AUG-2003
DEFINITION Brassica rapa cultivar Samjin reduced vernalization response 1
mRNA, complete cds.
ACCESSION AY356368
VERSION AY356368.1 GI:33943516
KEYWORDS
SOURCE Brassica rapa
ORGANISM Brassica rapa
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
AUTHORS Kwon,S.-J., Park,B.-S., Kim,S.-Y., Choi,H.-S., Lee,M.-C.,
Kim,J.-S., Lee,S.-I., Lim,K.-B., Kim,J.-A., Hong,K.-Y., Lee,M.-R.,
Jin,Y.-M., Kim,D. and Kim,H.-I.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2003) Genomics Division, National Institute of
Agricultural Biotechnology, Seodun-dong 225, Suwon 441-707,
Republic of Korea
FEATURES
source
1..990
/organism="Brassica rapa"
/mol_type="mRNA"
/cultivar="Samjin"
/db_xref="taxon:3711"
CDS
1..990
/note="VRN1"
/codon_start=1
/product="reduced vernalization response 1"
/protein_id="AAQ55453.1"
/db_xref="GI:33943517"
/translation="MPPFFHKLIFSSITQEKRLVPDFKVSFKDELSVAVALTVPD
GHVWVGLRKADNNKNIWFDQWQFVDVRSIRIGYLLIFRVEGNSAPSVCIYNLPQS
EINVHSTGLMDSASNNHFKRPLRLEDLEDEDAETLHTTASAIQSPFTGPVKPEBATP
TQTSKVPKGRKKXNADHPVEVNSAPRDDDPESRSKEYESASARKGTNABERRA
VNAATFEPTNPFVRVLRPSLYRIGCMILPSGFAEKLISGSIKVKQLGSKMPV
RCLYKAGRAKFSQGWYBFTVENNLGEGDVCVFELLRTDFVLKVTAYRVNYY"

ORIGIN

Query Match 65.5%; Score 673; DB 8; Length 990;
Best Local Similarity 82.5%; Pred. No. 2.4e-184; Mismatches 105; Indels 78; Gaps 5;
Matches 864; Conservative 0;
QY 1 ATGCCACGCGCTTCTTCCATAAGTTGATTTCTCATCCATATCCAGAAAAACGCTCTG 60
Db 1 ATGCCACGCGCTTCTTCCACAGCTGATTTCTCATCCACTATCCAGAAAAACGCTCTG 60
QY 61 AGGGTCCAGATAAGTTTGTGAGTAAATTCAGAGATGAGCTTTCGTTGCTGTGCACTC 120
Db 61 AGAGTTCAGATAAAGTTTGTGAGTAAATTCAGAGACGAGCTATCGTTGCACTC 120
QY 121 ACAGTACTGATGGTCATGTTTGGCGTGTAGGACTAAGGAAAGCTG-----ACAACAA 174
Db 121 ACAGTACCGGATGGTCATGTTTGGCGTGTAGGACTAAGGAAAGCTGCAACAAACAAG 180
QY 175 ATTTGGTTTCAAGATGTTTGGCAAGATTTGTGACCGTTTACTTCCATTCGCAATGGTTAT 234
Db 181 ATTTGGTTTCAAGATGTTTGGCAAGATTTGTGACCGTTTACTTCCATCCGATGGTTAC 240
QY 235 CTTTGTATTTTATAGATATGAAGAAACTCTGCTTCAGCGTCTACATTTTCAATTTATCC 294
Db 241 CTTTGTATCTTTAGATACGAAGCAACTCTGCTTCAGCGTGTGCACTTTTACAACTTACCA 300
QY 295 CACTCTGAGATCAATTAACCTTCCACCGGCTCTCATGGATTTCCGC-----TCACACACC 348
Db 301 CAGTCCGAGATTAACCTTCCACCGGCTCTCATGGACTCTGCACTACACAAACAACCAC 360
QY 349 TTTCAAAACGCGCGCTTTTGTTCGAAGACCTTGAAGATGAAGATGCCGAGGTCATCTTTCT 408
Db 361 TTTCAAAACGCGCGCTTTTGTTCGAAGACCTTGAAGATGAAGATGCTGAGACTCTGCACACA 420
QY 409 TCTTCTGTGTACCCACTCACCACTTCTCTGAGTCTACAGTACCAGCCAAACAAAGGTTAGCT 468


```
QY 760 TACAGAGTTGCATCATGATATCTCTCTTCGGTGTTCGAGAGTACCTAAGTGGATC 819
Db 724 TACAGAGTTGCATCATGATATCTCTCTTCGGTGTTCGAGAGTACCTAAGTGGATC 783
QY 820 TCCGGGTTTCATCAAGTCCAGCTTCGGGAGAACAAATGGCTGTTCGATGCTCTACAAA 879
Db 784 TCCGGGTTTCATCAAGTCCAGCTTCGGGAGAACAAATGGCTGTTCGATGCTCTACAAA 843
QY 880 GCCGGGAGAGCCAAATTCAGTCAAGGATGGTACGAATTCACCTACAGAACCACTTAGGA 939
Db 844 GCAGGGAGAGCTAAGTTCAGTCAAGGTTGTATGAGTTCACCTGGAGATAACCTAGGA 903
QY 940 GAAGGAGAGCTCTGTGTGTTCAGCTTCAGAACAGAGATTTCTGTTTGAAGTGACA 999
Db 904 GAAGGAGAGCTCTGTGTGTTCAGCTTCAGAACAGAGATTTCTGTTTGAAGTGAGC 963
QY 1000 GCCTTTTCGAGTCAACAGTACGTCGA 1026
Db 964 GCCTATCGAGTCAACGAGTACGTCGA 990

RESULT 8
AF289051
LOCUS
DEFINITION
Arabidopsis thaliana 5000 bp DNA linear PLN 05-NOV-2003
VRN1-Ler allele, complete cds.
ACCESSION
AF289051
VERSION
AF289051.1 GI:21734793
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosida II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 5000)
Levy,Y.Y., Mesnage,S., Wylne,J.S., Gendall,A.R. and Dean,C.
Multiple roles of Arabidopsis VRN1 in vernalization and flowering
time control
Science 297 (5579), 243-246 (2002)
JOURNAL
MEDLINE
22111275
PUBMED
22114624
REFERENCE
2 (bases 1 to 5000)
Levy,Y.Y., Gendall,A.R. and Dean,C.
VRN1, a gene required for response to vernalization
Unpublished
JOURNAL
3 (bases 1 to 5000)
Levy,Y.Y. and Dean,C.
Direct Submission
JOURNAL
Submitted (23-JUL-2000) Molecular Genetics, John Innes Centre,
Colney Lane, Norwich, Norfolk NR4 7UH, UK
FEATURES
Location/Qualifiers
1..5000
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="3"
/map="between mi339 and m560B2"
/ecotype="Landsberg erecta"
<1879..>4370
/gene="VRN1"
/allele="Ler"
/join(<1879..1938,3120..3567,3664..3732,3832..4031,
4122..>4370)
/gene="VRN1"
/allele="Ler"
/join(1879..1938,3120..3567,3664..3732,3832..4031,
4122..4370)
/gene="VRN1"
/allele="Ler"
/note="contains two VPI/ABI3-like B3 domains"
/codon_start=1
/product="reduced vernalization response 1"
/protein_id="AAM76972.1"
```

```
/db_xref="GI:21734794"
/translation="MPRPFFHKLIFSTIQEKLRLVRPDKFVSKFDLSVAVALTPVD
GHVWVGLRKADNKIWFQDMQWDFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSEI
NYHSTGLMDSAHNHFKARLPDLDEDEDAEIVPPSSVPSLPESPVPANKGVASSAI
OTLFTGVPVKAERTPTPKIPKXGRKKKNADPEEINSSAPRDDDDPENRSKFVESASR
KXIVTAEBERRAINAAKTFEPNPFVRVLRPSLYRGICIMYLPSPGFASKYLSIGSGF
IKVQAEQKQWVPVRCLYKAGRAKFSQGWYEFTELENNILGEGDVCFELLRTDRDFVLKVT
A
FRVNEYV"
ORIGIN
Query Match 63.6%; Score 653.4; DB 8; Length 5000;
Best Local Similarity 77.2%; Pred. No. 1.5e-178;
Matches 969; Conservative 0; Mismatches 1; Indels 285; Gaps 3;
QY 58 CTGAGGGTCCCAGATAAGTTTGTGAGTAAATCAAGATGAGCTTTCGGTGTGCTGTGCA 117
Db 3117 CAGAGGGTCCCAGATAAGTTTGTGAGTAAATCAAGATGAGCTTTCGGTGTGCTGTGCA 3176
QY 118 CTCACAGTACCTGATGGTCATGTTTGGCGGTGAGGCTAAGGAAAGCTGACAAACAAAT 177
Db 3177 CTCACAGTACCTGATGGTCATGTTTGGCGGTGAGGCTAAGGAAAGCTGACAAACAAAT 3236
QY 178 TGGTTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTCGCATTGGTTATCTT 237
Db 3237 TGGTTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTCGCATTGGTTATCTT 3296
QY 238 TTGATTTTATGATATGAAGGAAACTCTGCTTCAGCGTCTACATTTTCAATTTATCCCAC 297
Db 3297 TTGATTTTATGATATGAAGGAAACTCTGCTTCAGCGTCTACATTTTCAATTTATCCCAC 3356
QY 298 TCTGAGATCAATTACCATTCCACCGGTCTCATGGATTCGCTCACAAACCACTTCAACACG 357
Db 3357 TCTGAGATCAATTACCATTCCACCGGTCTCATGGATTCGCTCACAAACCACTTCAACACG 3416
QY 358 GCCCGTGTGTTGAAGACCTTGAAGATGAAGATGCGGAGTCACTCTTCTCTTCTCTG 417
Db 3417 GCCCGTGTGTTGAAGACCTTGAAGATGAAGATGCGGAGTCACTCTTCTCTTCTCTG 3476
QY 418 TACCATCACCATCTCTCGAGTCTACAGTACCAAGCAACAAAGGGTATGCTAGTTCAGCC 477
Db 3477 TACCATCACCATCTCTCGAGTCTACAGTACCAAGCAACAAAGGGTATGCTAGTTCAGCC 3536
QY 478 ATCCAAACCTTGTTCATCTGGACCAAGTTAA-----506
Db 3537 ATCCAAACCTTGTTCATCTGGACCAAGTTAAAGGTATATTTAATACCAACGATTCCTTT 3596
QY 507 -----506
Db 3597 ATCTATCGTGTATTAAGCGTCTTATCATTTTGTAGGTGTGATGTTGATATTTCTCTTA 3656
QY 507 -----AGCTGAAGAGCCAAACGCAACCCCAAAATACCTTAAAGAGAGAGGAGGAAG 561
Db 3657 TCTCCAGCTGAAGAGCCAAACGCAACCCCAAAATACCTTAAAGAGAGAGGAGGAAGAG 3716
QY 562 AAAAAATGCTGATCTCT-----576
Db 3717 AAAAAATGCTGATCTCTGGTAAAGCACTTTTCTCTTGAATGCTTTCAGACTCGTTTTCAGA 3776
QY 577 -----GAGGAA 582
Db 3777 GGATTCACAGATCTTCTCTCATGATACATATATCTTTTGTATATTTGCTTACAGAGGA 3836
QY 583 ATAAATCATCAGCTCCCGGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACAGAGAT 642
Db 3837 ATAAATCATCAGCTCCCGGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACAGAGAT 3896
QY 643 GCTTCTGCGAGAAAGAGAACCGTGTGCTGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 702
Db 3897 GCTTCTGCGAGAAAGAGAACCGTGTGCTGAGAGAAAGAGAGAGAGAGAGAGAGAGAG 3956
QY 703 AAAAAATGCGAACCAACAAACCTTTCTTCAGAGTGGTTCTCGGACCATCTTATCTATAC 762
Db 3957 AAAAAATGCGAACCAACAAACCTTTCTTCAGAGTGGTTCTCGGACCATCTTATCTATAC 4016
```


Qy	763	AGAGGTTGCATCAT	-----	776
Db	4017	AGAGGTTGCATCATCGTAATAAAAAAACAATCTTAGGAAGACTTAATCTTATCGGTGTC	TTT	4076
Qy	777	-----	-----GTATCTTCTCTCTGGG	792
Db	4077	CACGTATCTCTAAAGAAGCGCTTCTGTTCTCTCTCAACAGATATCTTCTCTCTGGG		4136
Qy	793	TTTGCTGAGAAGTACCTAAGTGGGATCTCCGGGTTTCATCAAAAGTCCAGCTTGGCGAGAAA		852
Db	4137	TTTGCTGAGAAGTACCTAAGTGGGATCTCCGGGTTTCATCAAAAGTCCAGCTTGGCGAGAAA		4196
Qy	853	CAATGGCCTGTTCCGATGTTCTTACAAAGCGGAGAGGCAAAATTCAGTCAAGGATGGTAC		912
Db	4197	CAATGGCCTGTTCCGATGTTCTTACAAAGCGGAGAGGCAAAATTCAGTCAAGGATGGTAC		4256
Qy	913	GAATTCACCTCTAGAGAACAACCTTAGAGAAGAGAGAGCTGTGCTGTTTGAGCTGCTCAGA		972
Db	4257	GAATTCACCTCTAGAGAACAACCTTAGAGAAGAGAGAGCTGTGCTGTTTGAGCTGCTCAGA		4316
Qy	973	ACCAGAGATTCCTGTTTGAAGTGACAGCGCTTTCGAGTCAACGAGTACGTCCTGAA		1027
Db	4317	ACCAGAGATTCCTGTTTGAAGTGACAGCGCTTTCGAGTCAACGAGTACGTCCTGAA		4371
RESULT 9				
AP000735/c				
LOCUS	AP000735	79186 bp	DNA linear	PLN 14-FEB-2004
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 3, TAC clone: K13E13.			
ACCESSION	AP000735 BA000014			
VERSION	AP000735.1 GI:6451856			
KEYWORDS				
SOURCE	Arabidopsis thaliana (thale cress)			
ORGANISM	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi			
REFERENCE	1			
AUTHORS	Kaneko, T., Katoh, T., Sato, S., Nakamura, A., Asamizu, E. and Tabata, S.			
TITLE	Structural analysis of Arabidopsis thaliana chromosome 3. II.			
	Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC			
	and BAC clones			
JOURNAL	DNA Res. 7 (3), 217-221 (2000)			
MEDLINE	20363099			
PUBLISHED	10907853			
REFERENCE	2 (bases 1 to 79186)			
AUTHORS	Kaneko, T., Kato, T., Sato, S., Nakamura, Y., Asamizu, E. and Tabata, S.			
TITLE	Direct Submission			
JOURNAL	Submitted (18-NOV-1999) Yasukazu Nakamura, Kazusa DNA Research			
	Institute, Department of Plant Gene Research; 1532-3, Yana,			
	Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,			
	Tel: 81-438-52-3935, Fax: 81-438-52-3934)			
	Address for correspondence: kaos@kazusa.or.jp			
COMMENT	For the latest information on annotation of this clone, please see			
	http://www.kazusa.or.jp/kaos/cgi-bin/agd/graph.cgi?c=K13E13			
	Genes with similarity to proteins in the databases are described in			
	'product' or 'note' qualifiers. Genes that have no significant			
	protein similarity are described as 'unknown protein'.			
	The software programs used to predict genes include: Grail			
	(Informatics Group, Oak Ridge National Laboratory,			
	http://combio.ornl.gov/Grail-1.3/)			
	GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),			
	NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of			
	Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and			
	SplicePredictor (Volker Brendel, Stanford University,			
	http://grainline.zozi.iastate.edu/cgi-bin/sp.cgi)			
	Genes encoding tRNAs are predicted by tRNAscan-SE			
	(Sean Eddy, Washington University School of Medicine, St. Louis,			
	http://genome.wustl.edu/eddy/tRNAcan-SE/).			
	This sequence may not be the entire insert of this clone. It may be			
	shorter because we remove overlaps between neighboring submissions.			
	The 5' clone is MCB22 and the 3' clone is MHP21.			

CDS	/evidence=not_experimental 19930..20860 /note="gb AAD43171.1 gene_id:K13E13.6 similar to unknown protein" /pseudo /codon_start=1 /evidence=not_experimental join(21250..21381,21635..21952,22001..22061,22130..22200) /note="unnamed protein product; gb AAC72857.1 gene_id:K13E13.7 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01692.1" /db_xref="GI:9280313" /translation="MYTONTKARVTSVSHRSQQDPSPVKKFKLVLPSTMKDRWM KIPPRVKLGSKLSEVVTLETPAGFKRSIKLKRIGEEIWFHEGSEFAEASIERGH FLFPEYKENSFRVLIENVSACETKPLDAVHIIDSDDDIIDIITKEFDEHKSKRS RDIEDKILHLDVMPMHVKEEBEEDKRVFRG" complement(22331..23749) /note="unnamed protein product; gb AAF15933.1 gene_id:K13E13.8 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01693.1" /db_xref="GI:9280314" /translation="MSSVFGPRFLSLLOONSKTLIQAKIHAQLVINGCHDSLFGK LIGHYCSKSTESSKLAHLVPPRGHDKFLFNLLCKSPEDSIRIPANYASKS LILYNERTFVLGACARSASSALVRGIRVHGMVKKGLGLYESELIIGTLLHFYAKN GLYLRARVDFEMPERTSVTNAMIGGYCSHKDKGNHAKMVLFRFSCGSGVRRP TDTTWVCVLSAISQTLGELSHVGYIEKLGTFPEVDVFIGTALVDMYKCGCLNAY FVVFELMKVNTFVTSMTALGALNGRNETNLLNMAESGIKPEIITFTSLISAYR HLSLREGIELEKSMKTRGVTPVIEHYCCIYDLGKAGRIQEAQVIFILLAPIKDAI LILSLNACSIYGETWGEIEIGKALLEIEREBEKLSGSECEYVALSNVLAHKGWVE VSKLRKEMERIKTRPGYSFV" 25146..26393 /note="unnamed protein product; gb AAC69119.1 gene_id:K13E13.9 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01694.1" /db_xref="GI:9280315" /translation="WTIPDLNDLVEILCRVPARNLKLRLSTSKRNRLFKDDRFEA REHMKAPKYLPLMTSEYRICPVSLNQGDPVSVLRELSPDPDYPHQDGRV FHCDDLGVNHVGRPRYGSKIYVWNPLGTQTRWIEAGRWKEYRFRVIGYCYQDE NNSCSKKIYKILCFYPNGQDTBIYELNYSIDRWRTIPDGLTPGWTLYISEQTYSNMG NLYLFASEKSPHLGVSLRFPDSTEKSLCVTLPYQRPRIELISAVRGGENISLL LQIDPESKTEIWTNKIDDTTKGAASVTKVLAFDLSPLQLFSEEVNFLDEDKV AVCCERLPEQEHRYOCRRVKITDKIYILGEDNKVDVSGSEGEATDSLGIQVI LNVAPSLVIOEAGGGKTRGDD" complement(join(27709..27957,28048..28247,28347..28409, 28512..28692)) /note="unnamed protein product; gb AAD43153.1 gene_id:K13E13.10 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01695.1" /db_xref="GI:9280316" /translation="MDSAHNFKRLRFLDEDEDAEIVFPSSVYPSLPPESTVPANK GYVASAIOTLFTGPVKPEPTPKIKKRGKKKNADPEINSSAPRDDDDPENRSKFE SASARKTVAERERAINAKTPTNPFRRVLRPSYLYRGCIWYLPSPGAEKYLS GISGFIKVLAKQMPVRLYKAGRAKFSQGWYFTLENNLGEEDVCFPELLRTDFV LKVTAFRVNEYV" complement(join(32300..32554,32669..32996,33144..33619)) /note="gene_id:K13E13.11" /codon_start=1 /evidence=not_experimental /product="oxylase-like protein" /protein_id="BAB01696.1"	/db_xref="GI:9280317" /translation="MGELDEAFIQAEHRPNTHLTNSGDFIFSDIEPTIDLSLEDTH HDKTAIAKEIAEACRGGFFQVINHGLPSALHRVVKTAAEFNLFTTEBRKVKVROEV NPMGVHDBERTKNVRDWKEIFDFELQDSTIVPASPEDETELKLTNOMPONPSHRE VCOBYAREVEKELAPRLLELVSISLGLPGRLTGFNEQTSFLRFNHPPCPNPELALG VGRHKDGGALTIVLAQDSVGLQVSRSDGMIPVKFISDALIINMGNCIQVWNTDFTW SABHRVNTSNERFSIFPFFPSHEANIEPLEEELISEENPPCYKKNWKGKFFVSRNR SDPKKLEVENIQIDHFKA" complement(join(34938..35198,35304..35631,36034..36494))	Query Match 63.6%; Score 653.4; DB 8; Length 79186; Best Local Similarity 77.2%; Pred. No. 2.1e-178; Indels 285; Gaps 3; Matches 969; Conservative 0; Mismatches 1;	QY 58 CTGAGGGTCCAGATGAAGTTTGTGAGTAAATTCAGAGATGAGCTTTCGGTTCGTGTGCA 117 DB 28962 CAGAGGGTCCAGATGAAGTTTGTGAGTAAATTCAGAGATGAGCTTTCGGTTCGTGTGCA 28903 QY 118 CTCACAGTACCTGATGGTCAATGTTTGGCGGTGTAGGACTAAGGAAAGCTGACAAATAATT 177 DB 28902 CTCACAGTACCTGATGGTCAATGTTTGGCGGTGTAGGACTAAGGAAAGCTGACAAATAATT 28843 QY 178 TGGTTTCAAGATGGTTGGCAAGAGTTTGTGACCGGTACTCCATTGCGATTGGTTATCTT 237 DB 28842 TGGTTTCAAGATGGTTGGCAAGAGTTTGTGACCGGTACTCCATTGCGATTGGTTATCTT 28783 QY 238 TTGATTTTATAGATATGAAGAAATCTCGCTTCAGCGTCTACATTTTCAATTTATCCAC 297 DB 28782 TTGATTTTATAGATATGAAGAAATCTCGCTTCAGCGTCTACATTTTCAATTTATCCAC 28723 QY 298 TCTGAGATCAATTAACATTCACCGGTCTCATGGATTCCGCTCACAAACACTTCAAACGC 357 DB 28722 TCTGAGATCAATTAACATTCACCGGTCTCATGGATTCCGCTCACAAACACTTCAAACGC 28663 QY 358 GCCCGTTGTTTGAAGACCTTGAAGATGAAGTCCGAGGTCACTCTTCTCTTCTTCTG 417 DB 28662 GCCCGTTGTTTGAAGACCTTGAAGATGAAGTCCGAGGTCACTCTTCTCTTCTTCTG 28603 QY 418 TACCATTACCACTTCTCTGAGTCTACAGTACCAAGCAACAAAGGGTATGCTAGTTCAGCC 477 DB 28602 TACCATTACCACTTCTCTGAGTCTACAGTACCAAGCAACAAAGGGTATGCTAGTTCAGCC 28543 QY 478 ATCCAAACCTTCTCTCACTGACCAAGTTAA-----506 DB 28542 ATCCAAACCTTCTCTCACTGACCAAGTTAAAGGTATATTTATAACCAACTGATTCCTTT 28483 QY 507 -----506 DB 28482 ATCTATCGCTGATTACGGCTCTTATCATTTCTTTGAGGTGTGATATTTCTCTTA 28423 QY 507 -----AGCTGAAGAGCAACGCCAACCCCAAAAATACCTAAAAAGAGAGAGGAGAGAAG 561 DB 28422 TCTCCAGCTGAAGAGCAACGCCAACCCCAAAAATACCTAAAAAGAGAGAGGAGAGAAG 28363 QY 562 AAAAAATGCTGATCTCT-----576 DB 28362 AAAAAATGCTGATCTCTGGTAAAGCACTTTCTCTTGAATGCTTCAGACTCGTTTTCAGA 28303 QY 577 -----GAGGAA 582 DB 28302 GGATTCACAGATCTTCTCTCATGATACATATATCTCTTTTGATATTTGCTCTTACAGAGAA 28243 QY 583 ATAAACTCATAGCTCCGCGAGATGATGATCCAGAGAACCGTTTCAAAGTTCTTACGAGAGT 642 DB 28242 ATAAACTCATAGCTCCGCGAGATGATGATCCAGAGAACCGTTTCAAAGTTCTTACGAGAGT 28183 QY 643 GCTTCTCGAGAAAGAGAACCGGTGCTGAGTGCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 702 DB 28182 GCTTCTCGAGAAAGAGAACCGGTGCTGAGTGCAGAAAGAGAGAGAGAGAGAGAGAGAGAG 28123 QY 703 AAAAACTGGAACCAACAAACCCCTTTCTTTCAGAGTGGTTCTCGGACCACTCTATCTATATAC 762 DB 28122 AAAAACTGGAACCAACAAACCCCTTTCTTTCAGAGTGGTTCTCGGACCACTCTATCTATATAC 28063
-----	---	---	---	--

QY 763 AGAGTTGCATCAT----- 776
DB 28062 AGAGTTGCATCATGTAATAAAAAACATCTTAGGAAGACTTAATCTTATCGGTGCTT 28003
QY 777 -----GTATCTTCTCTCTGGG 792
DB 28002 CACTGATCTCTAAAAGAAGCCTTCTGTTTCTGTTTCTCTCAACAGTATCTTCTCTGGG 27943
QY 793 TTTCGTGAGAGTACCTAAGTGGGATCTCCGGGTTCATCAAGTCCAGCTTCCGGAGAAA 852
DB 27942 TTTCGTGAGAGTACCTAAGTGGGATCTCCGGGTTCATCAAGTCCAGCTTCCGGAGAAA 27883
QY 853 CAATGGCTGTTCGATGCTCTCAAAAGCCGGAGAGCCAAATTCAGTCAAGGATGCTAC 912
DB 27882 CAATGGCTGTTCGATGCTCTCAAAAGCCGGAGAGCCAAATTCAGTCAAGGATGCTAC 27823
QY 913 GAATTCATCTAGAGAACAACTTAGGAGAGGAGAGCGTGTGTGTTGAGCTGCTCAGA 972
DB 27822 GAATTCATCTAGAGAACAACTTAGGAGAGGAGAGCGTGTGTGTTGAGCTGCTCAGA 27763
QY 973 ACCAGAGATTTGGTTTGAAGTACAGAGCTTTTCGAGTCAACGAGTACGCTGAAA 1027
DB 27762 ACCAGAGATTTGGTTTGAAGTACAGAGCTTTTCGAGTCAACGAGTACGCTGAAA 27708

RESULT 10
AX101027
LOCUS AX101027 5000 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 1 from Patent WO0121822.
ACCESSION AX101027
VERSION AX101027.1 GI:13619883
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
Dean, C. and Levy, V. Y.
METHODS Methods and means for modification of plant flowering
characteristics
TITLE Patent: WO 0121822-A 1 29-MAR-2001;
JOURNAL Plant Bioscience Limited (GB)
FEATURES
source Location/Qualifiers
1..5000
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 63.6%; Score 653; DB 6; Length 5000;
Best Local Similarity 77.1%; Pred. No. 1.9e-178;
Matches 968; Conservative 1; Mismatches 1; Indels 285; Gaps 3;
QY 58 CTGAGGGTCCAGATAGTTGTGAGTAAATCAAGATGAGCTTTCCGTTGCTGTGCA 117
DB 3117 CAGAGGGTCCAGATAGTTGTGAGTAAATCAAGATGAGCTTTCCGTTGCTGTGCA 3176
QY 118 CTCACAGTACCTGATGCTATGTTGGCGTGTAGGACTAAGAAAGCTGACAAATAAT 177
DB 3177 CTCACAGTACCTGATGCTATGTTGGCGTGTAGGACTAAGAAAGCTGACAAATAAT 3236
QY 178 TGGTTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTCGATTCGTTATCTT 237
DB 3237 TGGTTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTCGATTCGTTATCTT 3296
QY 238 TTGATTTTATAGATAGAGAAACTCTGCTTTCAGCGTCTACATTTTCAATTTATCCAC 297
DB 3297 TTGATTTTATAGATAGAGAAACTCTGCTTTCAGCGTCTACATTTTCAATTTATCCAC 3356
QY 298 TCTGAGATCAATACCATTCACCGGCTCTCATGGATTCGGTTCACCAACCACTTCAAGCG 357

DB 3357 TCTGAGATCAATACCATTCACCGGTCTCATGGATTCCGCTCACACCACTTCAACGC 3416
QY 358 GCCCGTTTGTGAGAGACCTTGAAGATGAGATGCGAGGTCACTTCTTCTTCTGTG 417
DB 3417 GCCCGTTTGTGAGAGACCTTGAAGATGAGATGCGAGGTCACTTCTTCTTCTGTG 3476
QY 418 TACCCATCAACCACTTCTGAGTCTACAGTACAGCCCAAAAGGATAGCTAGTTGAGCC 477
DB 3477 TACCCATCAACCACTTCTGAGTCTACAGTACAGCCCAAAAGGATAGCTAGTTGAGCC 3536
QY 478 ATCCAAACCTTCTTCACTGGACCACTAA----- 506
DB 3537 ATCCAAACCTTCTTCACTGGACCACTTAAAGGTGATATTATAACCAACTGATTCCTTT 3596
QY 507 ----- 506
DB 3597 ATCTATCGCTGATTACGGGTCTTATCATTTCTTTGAGGTGATGCTTGATATTTTCTT 3656
QY 507 -----AGCTGAAGAGCCCAACGCCAACCCCAAAATACCTTAAAGAGAGGAGGAGAG 561
DB 3657 TCTCAGCTGAGAGAGCAACGCCAACCCCAAAATACCTTAAAGAGAGGAGGAGAGAG 3716
QY 562 AAAAAATGCTGATCTT----- 576
DB 3717 AAAAAATGCTGATCTTGGTAAGCACTTTTCTCTTTGAAATGCTTCAGACTCGTTTTCAGA 3776
QY 577 -----GAGAA 582
DB 3777 GGATTCACAGATTTCTTCTCATGATACATATATCTTTTGATATTGCTTCTTACAGAGAA 3836
QY 583 ATAAACTCATCAGCTCCGGAGATGATCCAGAGACCGTTTCAAAAGTTCTACGAGAGT 642
DB 3837 ATAAACTCATCAGCTCCGGAGATGATCCAGAGACCGTTTCAAAAGTTCTACGAGAGT 3896
QY 643 GCTTCTCGAGAGAGAGAACCGTCACTGACAGAGAGAGAGAGAGAGCCATCAATGACGCC 702
DB 3897 GCTTCTCGAGAGAGAGAACCGTCACTGACAGAGAGAGAGAGAGAGCCATCAATGACGCC 3956
QY 703 AAAAGTTGCAACCAACCAACCTTCTTCAAGAGTGGTTCTGCGACCACTCTATCTATAC 762
DB 3957 AAAAGTTGCAACCAACCAACCTTCTTCAAGAGTGGTTCTGCGACCACTCTATCTATAC 4016
QY 763 AGAGTTGCATCAT----- 776
DB 4017 AGAGTTGCATCATGTAATAAAAAACATCTTAGGAAGACTTAATCTTATCGGTGCTT 4076
QY 777 -----GTATCTTCTTCTGGG 792
DB 4077 CACTGATCTCTAAAAGAAGCCTTCTGTTTCTCTCAACAGATATCTTCTTCTGGG 4136
QY 793 TTTCGTGAGAGTACCTAAGTGGGATCTCCGGGTTCATCAAGTCCAGCTTCCGGAGAAA 852
DB 4137 TTTCGTGAGAGTACCTAAGTGGGATCTCCGGGTTCATCAAGTCCAGCTTCCGGAGAAA 4196
QY 853 CAATGGCTGTTCGATGCTCTCAAAAGCCGGAGAGCCAAATTCAGTCAAGGATGCTAC 912
DB 4197 CAATGGCTGTTCGATGCTCTCAAAAGCCGGAGAGCCAAATTCAGTCAAGGATGCTAC 4256
QY 913 GAATTCATCTAGAGAACAACTTAGGAGAGGAGAGCGTGTGTGTTGAGCTGCTCAGA 972
DB 4257 GAATTCATCTAGAGAACAACTTAGGAGAGGAGAGCGTGTGTGTTGAGCTGCTCAGA 4316
QY 973 ACCAGAGATTTGGTTTGAAGTACAGAGCTTTTCGAGTCAACGAGTACGCTGAAA 1027
DB 4317 ACCAGAGATTTGGTTTGAAGTACAGAGCTTTTCGAGTCAACGAGTACGCTGAAA 4371

RESULT 11
AY042868
LOCUS AY042868 1146 bp mRNA linear PLN 04-JUL-2001
DEFINITION Arabidopsis thaliana Unknown protein (F13P21.8) mRNA, complete cds.
ACCESSION AY042868
VERSION AY042868.1 GI:14596160

QY 513 AGAGCCAAAGCCCAACCCCAAAATACCTAAAGAGAGAGGAGGAAGAAAGAAAATGCTGA 572
Db 28625 AGA-----AACGCCCAAGGTACTTAAGAGAGAGAGGAAGGAAGAAATCCTAA 28675
QY 573 TC----- 574
Db 28676 TCCCGTAAGTATCAGAGACAGACTCTTAATACGTATTAACTTTTGTTCAGAGGAG 28735
QY 575 -----CTGAGAAATAAATCATCATCAGCTCCGCGAGATG 607
Db 28736 TCAAAAGATTATTTGTTTGGTATGTGAGAGAGAGTAATCTTTCACTCCCGTGGAG 28795
QY 608 ATGATCCAGAGAACCGTTCACAAAGTCTTACGAGAGTGCTTCTGCGAGAAAGAACCGTGA 667
Db 28796 ATGATCTCAGAGAACCGCTCAAAAGTCTTACGAGAGTGCTTCTGCTAGAAAGAGAACTGTAA 28855
QY 668 CTGCAAGAGAGAGAGAGGAGCCATCAATGAGCCCAAAACGTTCCGAAACCAACCCCTT 727
Db 28856 CTGCAAGAGAGAGAGAGGAGCCGTCATGCGAGCCCAAAACATTCGAAACCAAAATCCTT 28915
QY 728 TCTTCAGAGTGCTTCTGCGACCATCTATCTATACAGAGGTTGCAATCAT----- 776
Db 28916 ACTTTAGAGTTGTTCTGCGACCATCATATCTATACAGAGGTTGCATCATGTTAAACAATT 28975
QY 777 ----- 776
Db 28976 TTTAAAGATTGTAAATTAAGGTTCAAGATTAGGGTTTAACATGACCTGTCTTCTGCTGACTTC 29035
QY 777 -----GTATCTCTCTCTGCTGGTTTCTGAGAGTACC 808
Db 29036 TTGCTTTCTGTTTTTTGTTTCAACAGTACTTCCCATCTGCGGTTTCTGAGAAATACC 29095
QY 809 TAAAGTGGATCTCCGGGTTTCATCAAAAGTCCAGCTTTCGCGAGAGAAACAAATGGCTGTTCGAT 868
Db 29096 TAAAGTGGATATCTGTTTTCATCAGCTCCAGCTCGGTGAGAAACAAATGGCCAGTGAGT 29155
QY 869 GTCTCTACAAGCCGGGAGAGCAAAATCAGTCAAGGATGGTACGAATTCACCTCTAGAGA 928
Db 29156 GCCTCTACAAGCAGGAGAGCTAAGTTTAGCCAAGGATGGTATGAGTTCACTCTCAGAGA 29215
QY 929 ACAACTTAGGAGAGGAGAGCTGCTGTTGTTGAGCTGCTCAGACCCAGAGATTCGTTT 988
Db 29216 ACAATATAGGCGAAGGAGATGATGTGTTGAGCTACTCAGAACTCGGGAATTCGTTTC 29275
QY 989 TGAAAGTGACAGCCTTTTCAGTCAACGAGTACGTCTGA 1026
Db 29276 TCGAAGTCAACCGCCTTTGCTGTCATGATGATGTCTGA 29313

RESULT 14
AP006145
LOCUS AP006145 78589 bp DNA linear PLN 22-JUL-2003
DEFINITION Lotus corniculatus var. japonicus genomic DNA, chromosome 2,
clone:LjT03H18, TM0262, complete sequence.
ACCESSION AP006145
VERSION AP006145.1 GI:29122784
KEYWORDS HTG.
SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
Lotus.
REFERENCE 1
AUTHORS Kaneko,T., Asamizu,E., Kato,T., Sato,S., Nakamura,Y. and Tabata,S.
TITLE Structural analysis of a Lotus japonicus genome. III. Sequence
features and mapping of sixty-two TAC clones which cover the 6.7 Mb
regions of the genome
JOURNAL DNA Res. 10 (1), 27-33 (2003)
MEDLINE 22579290
PUBMED 12693552
REFERENCE 2 (bases 1 to 78589)
AUTHORS Sato,S.

TITLE Direct Submission
JOURNAL Submitted (12-DEC-2002) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba 292-0818, Japan (E-mail:ssato@kazusa.or.jp,
URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),
Fax:81-438-52-3934)
FEATURES
source Location/Qualifiers
1. 78589
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="2"
/clone="LjT03H18"
/clone_lib="LjT library"
/note="TAC clone:TM0262-synonym: Lotus japonicus"
ORIGIN
Query Match 13.7%; Score 140.4; DB 8; Length 78589;
Best Local Similarity 66.6%; Pred. No. 4.4e-29;
Matches 201; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 21 TAAAGTTGATTTTCTCATCCACTATCCAGAAAACGCTCGAGGGTCCAGATAAGTTTGT 80
Db 15994 TACGAATATTTGAAGCAACTTGTATTTTGAATTCAGGGGATTCGGACAATTTCT 16053
QY 81 GAGTAAATTCAGGATGAGCTTTCCGTTTCTGCTGCTCACTCACAGTACCTGATGTCATGT 140
Db 16054 GAGGAATATGCGGTGAGCTTTCAACAATTTGCTACCTCACTGTCAGATGGTACTGT 16113
QY 141 TTGCGGTGTAGACTAAGGAAGCTGACAAACAAATTTGGTTTCAAGATGGTTGGCAAGA 200
Db 16114 TTGCGGTGTAGGGTTAAAGAGGCGCGACACAGAATTTGGTTGCTGATGGTTGGCAAGA 16173
QY 201 GTTTGTTGACCGTTACTCCATTCGCTTCCATTTGTTTATCTTTTGTATTTAGATGAGGAAA 260
Db 16174 ATTTGTTCAACGCTACACCAATTCGCAATTCGATGATTTAGTATTCATGATGAGGAAA 16233
QY 261 CTCTGCTTACGCGTCTACATTTTCAATTTATCCCACTCTCAGATCAATTACCATTCCAC 320
Db 16234 TCCATCTTCATTTTCATATTTTCAATATGAGCACTTCGAGGTGAATTATCAGTCTGC 16293
QY 321 CG 322
Db 16294 AG 16295
RESULT 15
AC137825/c
LOCUS AC137825 243777 bp DNA linear HTG 13-AUG-2004
DEFINITION Medicago truncatula clone mth2-23j1, WORKING DRAFT SEQUENCE, 17
unordered pieces.
ACCESSION AC137825
VERSION AC137825.22 GI:50897368
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 243777)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE Medicago truncatula BAC Clone mth2-23j1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 243777)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Farrington Oval, Room 208, Norman,
OK 73019, USA

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FLI CDNA.
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1146)
 Nguyen, M., Karlin-Neumann, G., Southwick, A., Lam, B., Miranda, M.,
 Palm, C.J., Bowser, L., Jones, T., Ban, J., Carninci, P., Chen, H.,
 Cheuk, R., Chung, M.K., Hayashizaki, Y., Ishida, J., Kamiya, A.,
 Kawai, J., Kim, C., Lin, J., Liu, S.X., Narusaka, M., Pham, P.K.,
 Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K.,
 Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.
 Direct Submission
 Submitted (25-JUN-2001) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
 Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PDEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Nguyen, M.,
 Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C.J.,
 Bowser, L., Jones, T., Ban, J., Chen, H., Cheuk, R., Chung, M.K.,
 Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H., Shinn, P.,
 Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Nguyen, M., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed
 equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W.
 (SSP/Stanford) contributed equally to this work as PIs.

FEATURES
 source
 Location/Qualifiers
 1..1146
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="RAFL04-16-P22"
 /ecotype="Columbia"
 /note="This clone is in pBluescript"
 1..1146
 /gene="F13F21.8"
 177..857
 /codon_start=1
 /product="Unknown protein"
 /protein_id="AAK68808.1"
 /db_xref="GI:14596161"
 /translation="WQMDSAQNFNRKARLPEDPELKDQKVIYPISTEPVNGY
 GSTDQSPFKESKABETPKVLKGRKKKNPPEVNSVTPGGDSENRSKEYESASA
 RKTVAERAEVNAAKTEPTNPYFVRLPSYLYRGICIMYLPSPGFAEKLISIG
 FIKLQIGERQWVRCLYKAGRAFSGQWYEFTLENNIGEDVCVFELLRLTRDFVLEIT
 AFRVNEYV"

ORIGIN
 Query Match 43.1%; Score 442.6; DB 8; Length 1146;
 Best Local Similarity 73.6%; Pred. No. 2.8e-117;
 Matches 647; Conservative 0; Mismatches 194; Indels 38; Gaps 5;
 QY 151 GGACTAAGGAAGCTGACACAAATTTGGTTTCAAGATGTTGGCAGAGTTTGTGAC 210
 DB 14 GGACTAAGGAAGCCCAACAAATCTGTTTCAAGACGGTGGCAGAGTTTGTGAC 73
 QY 211 CGTTACTCCATTCGATTCGTTATCTTTGATTTTATGATATGAGGAACCTCGCTTC 270
 DB 74 CGTTCTCATTCGGATTCGT-----TTCAGATACAAAGTTACAGTTACATTTCA 129
 QY 271 AGCGTCTAATTTTCAATTTATCCCACTCTGAGATCAATATCCACCGGTCTCATG 330

Db 130 ATCTCTCCACACTCTGAGATCAACACC-----ATTCTAGTAGTGAAGCTCTTATGCAATG 185
 QY 331 GATTTCGGCTCACAAACACTT---CAAACGCGCCGTTTGTTCGAAGACCTTGAAGATGAA 387
 Db 186 GATTTCGGCAGAGATCAGTTCAACAAACGTGCTCGATTTGTTGAAGATCTGAACTCAA 245
 QY 388 GATGCCGAGGTCACTTTCTTCTCTGTGTATCCCATCACTTCTCTGAGTCTACAGTA 447
 Db 246 GATGCTAAGGTCAATTATC-----CATCGATCCCTGAATCTACTGAA 287
 QY 448 CAGCCCAACAAAGGTATCTAGTTTTCAGCATCAACACTTGTTCAGTGGACAGTTAAA 507
 Db 288 CAGTGAATAAAGGTATAGCGGTCTTACAGACATCCAAAGCTTTTCAAGAATCTAAA 347
 QY 508 GCTGAAGAGCCAAACGCCAACCCCAAAATACCTTAAAGAGAGAGGAGGAGAGAGAAAAT 567
 Db 348 GCTGAAGA-----AACGCCAAGGTACTTTAGAGAGAGAGAGAGAGAGAGAT 398
 QY 568 GCTGATCTTGAGGAAATAAACTCATCAGCTCCGCGAGATGATGATCCAGAGAACCGTTCA 627
 Db 399 CCTAATCCCGAGAGAGTAACACTTTCACCTCCCGGTGGAGATGACTCAGAGAACCGCTCA 458
 QY 628 AGTTCTACAGAGTCTTCTCGGAGAAAGAGAACCGTCTACGACAGAGAGAGAGAGA 687
 Db 459 AAGTTCTACGAGAGTCTTCTGCTAGAAAGAGAGAACTGTAACTGCAGAGGAGAGAGAGA 518
 QY 688 GCATCAATGACGACCAAAAGCTTTCGAACCAACAAACCCCTTCTTCAGAGTGGTTTCGCGA 747
 Db 519 GCGTCAATGACGACCAAAACATTCGAACCAACAAATCCTTACTTTAGAGTTGTTTCGCGA 578
 QY 748 CCATCTATCTATACAGAGTTTGCATCATGTATCTTCTCTGGGTTTCTCGAGAAAGTAC 807
 Db 579 CCATCATATCTATACAGAGTTTGCATCATGTATCTTGCATCTGGGTTTCTCGAGAAATAC 638
 QY 808 CTAAGTGGATCTCCGGGTTTCATCAAGTCCAGTCTCGGAGAGAAACATAGTCTGTTCGA 867
 Db 639 CTAAGTGGATATCTGGTTTCATCAAGCTCCAGCTCGGTGAGAGAAACATAGGCGAGG 698
 QY 868 TGCTCTCAAAAGCCGGAGAGCCAAATTCAGTCAAGGATGTTGACCAATTCATCTAGAG 927
 Db 699 TGCTCTCAAAAGCCGGAGAGCTTAAGTTTACCAAGGATGTTGATGATTCACATTCGAG 758
 QY 928 AACAACTTAGGAGAGGAGACCTCTGTGTGTTTGTAGTCTCTCAGAACCAAGATTTTCGTT 987
 Db 759 AACATATAGGCGAGGAGATCTATGTGTGTTTGTAGTCTCAGAACTCGGATTTTCGTT 818
 QY 988 TTGAAAGTACAGCCCTTTTCGAGTCAACGAGTACGCTGA 1026
 Db 819 CTCGAAGTCAACCGCTTTTCGTTCAATGATGATGTGTA 857

RESULT 12
 AX505578 681 bp DNA linear PAT 27-SBP-2002
 LOCUS Sequence 273 from Patent WO0216655.
 DEFINITION AX505578
 ACCESSION AX505578
 VERSION AX505578.1 GI:23386815
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1
 REFERENCE
 AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
 TITLE Stress-regulated genes of plants, transgenic plants containing
 same, and methods of use
 JOURNAL Patent: WO 0216655-A 273 28-FEB-2002;
 The Scripps Research Institute (US); Syngenta Participations AG
 (CH)
 FEATURES
 Location/Qualifiers
 1..681
 /organism="Arabidopsis thaliana"

/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Query Match 38.7%; Score 397.6; DB 6; Length 681;
Best Local Similarity 76.6%; Pred. No. 3.2e-104;
Matches 538; Conservative 0; Mismatches 134; Indels 30; Gaps 3;

QY 328 ATGGATTCGGCTCACACACACTT---CAAAGCGCGCCCTTTGTTTGAAGACCTTGAAGAT 384
Db ATGGATTCGGCAGAGATCAGTTCACAAACGTCGCTCGATTGTTTGAAGATCTCGAATC 66

QY 385 GAAGATCGCGAGGTCACTTTCTTCTCTGTGTATCCCATCCACTTCTGTGACTAC 444
Db AAAGATGCTAAGGTCAATTTATCCATCGAAC-----CCTGAATCTACT 108

QY 445 GTACCGACCAACAAGGGTATCTAGTTCAGCCATCCAAACCTGTTCTCACTGGACCACTT 504
Db GAACCATGTAATAAGGTTATGCGGTTCTACAGCCATCCAAAGCTTTTCAAGAATCT 168

QY 505 AAAGCTGAAGAGCCACGCGCAACCCCAAAAATACCTAAAGAGAGGAGGAGGAAGAAA 564
Db AAAGCTGAAGA-----AAGGCCAAGGTACTTAAGAGAGAGGAGGAAGGAAGAG 219

QY 565 ATGCTGATCTGAGGAATAAATACTCATCAGCTCCGCGAGATGATGATCCAGAGAACGCT 624
Db AATCTCTAATCCCGAGGAAGTAACCTTCACTCCCGTGGAGATGACTCAGAGAACCGC 279

QY 625 TCAAGTTCTACGAGAGTCTTCTGCGAAGAGAACCGTGACCTGCGAAGAAAGAGAG 684
Db TCAAGTTCTACGAGAGTCTTCTGCGAAGAGAACCTGTAAGAGAGACTGTAATCGCAGAGAAAGAG 339

QY 685 AGAGCCATCAATGACGCCAAAAGCTTCGAAACCAACCAACCTTTCTTCAAGTGTCTG 744
Db AGAGCCGCTCAATGACGCAAAAACATTCGAACCAACCAACCTTTCTTCAAGTGTCTG 399

QY 745 CGACCATCTCTATACAGAGGTGTCATCATGATCTTCTTCTGCGTTCGTCGAGAG 804
Db CGACCATCATATCTATACAGAGGTGTCATCATGATCTTCTTCTGCGTTCGTCGAGAA 459

QY 805 TACCTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGCTTTCGCGAAGAAAACATGGCTGT 864
Db TACCTAAGTGGGATATCTGTTTTCATCAAGTCCAGCTCCGTCGCGAAGAAAACATGGCCAGTG 519

QY 865 CGATGTTCTCAAAAGCGGAGAGCCAAATTCAGTCAGAGGTATGATGATTCATCTCA 924
Db AGGTGCTCTTACAAAGCAGGAGGAGCTAAGTTTAGCCAGAGGTGATGATGATTCACATC 579

QY 925 GAGAACCACTTAGGAGAGGAGAGCTGTGTGTGTTTGAAGTCTGCTCAGAACGAGATTC 984
Db GAGAACCAATATAGCGAGAGGAGATGATGTGTGTTTGAAGTCTGCTCAGAACCTCGGATTC 639

QY 985 GTTTTGAAGTGAAGCTTTTCAGTCAACGAGTACGTCGTA 1026
Db GTTCTCGAAGTCAACCGCTTTTCGTGTCATGAGTATGTGTA 681

RESULT 13

AC007504 125021 bp DNA linear PLN 30-OCT-2002
LOCUS Arabidopsis thaliana chromosome I BAC F13P21 genomic sequence,
DEFINITION complete sequence.
AC007504
AC007504.3 GI:5430745
HTG.
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 125021)
REFERENCE
AUTHORS Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,

JOURNAL
REFERENCE
AUTHORS

Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.S.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Unpublished
2 (bases 1 to 125021)
Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (06-MAY-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

TITLE
JOURNAL

REFERENCE
AUTHORS

3 (bases 1 to 125021)
Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (08-JUN-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

TITLE
JOURNAL

REFERENCE
AUTHORS

4 (bases 1 to 125021)
Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (09-JUL-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

TITLE
JOURNAL

COMMENT

On Jul 9, 1999 this sequence version replaced gi:5019265.
e-mail for correspondence: arabesequence.stanford.edu
Genes with similarity to proteins in the databases are described
as 'putative', '-like' or 'similar to'. Genes that have EST
similarity but no significant protein similarity are described as
'unknown proteins'. Genes that are annotated based only on gene
prediction software are described as 'hypothetical proteins'.
The software programs used to predict genes include: Graal
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/section/index.html), GENSCAN (Chris Burge,
& A.Salamov, Sanger Centre, http://genomic.sanger.ac.uk/),
NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/NetPlantGene.html) and
eMotif (Nevill-Manning, C.G. Wu, T.D. & Brutlag, D.L.,
http://motif.stanford.edu/projects.html).
Location/Qualifiers
1..125021
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/map="mi208"
/clone="F13P21"
/complement(4079..5020)
/gene="F13P21.1"
complement(join(4079..4465,4631..5020))
/gene="F13P21.1"
/note="Hypothetical Protein"
/codon_start=1
/protein_id="AAD43146.1"
/translation="MTNLISGSSYPTANLYFMQVWKIECWLRHAFSVDETICQWEI
MTLKEKYWEYSDTLAAVDFPRLKLCLEYCFSTDLRLTSKRLAHVSKYIKLF
KAYKRPSSITSSSOVETLEEDIPAGYGVFAVFSQKVGSGKSLDIYLGPTLDM
AFRHFVLAHWKDNCRKEKLSMCAVDLSIPTTVASESSFSIGSVLSKYRSLLP
ENIQALICTRNWLRGFPKGESEEEVEEKEEKEEKA"
6527..10142
/gene="F13P21.2"

FEATURES
source

gene

CDS

gene

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2004, 09:49:40 ; Search time 601 Seconds
(without alignments)
8970.309 Million cell updates/sec

Title: US-10-088-187a-10_COPY_269_1295
Perfect score: 1027
Sequence: 1 atgcacgcccctttctcca.....agtcacgagtagctgtctgaa 1027

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1027	100.0	1495	4	Aaf62446 A thalian
2	1026	99.9	1026	12	Adh73922 Thale cre
3	1025.4	99.8	1510	3	Aac39629 Arabidops
4	1016	98.9	1509	3	Aac31591 Arabidops
5	934.4	91.0	1347	3	Aac51587 Arabidops
6	653	63.6	5000	4	Aaf62416 A thalian
7	548.4	53.4	753	3	Aac37903 Arabidops
8	441	42.9	1394	3	Aac54934 Arabidops
9	441	42.9	1396	3	Aac53657 Arabidops
10	439.4	42.8	1344	3	Aac37617 Arabidops
11	397.6	38.7	681	6	Abz12468 Arabidops
12	295.4	28.8	426	8	Abx62122 Arabidops
13	101	9.8	2000	6	Abz15168 Arabidops
14	66.4	6.5	796	3	Aac35159 Arabidops
15	63	6.1	411	8	Abx62454 Arabidops
16	61.2	6.0	630	3	Aac51585 Arabidops
17	61.2	6.0	786	3	Aac51592 Arabidops
18	57	5.6	102	12	Adp92625 Cotton ex
19	42.6	4.1	2000	8	Ada71938 Rice gene
20	42.6	4.1	4323	6	Abz14722 Arabidops
21	39.2	3.8	1211	3	Aac36162 Arabidops

C 22	38.8	3.8	2000	8	ADA71938	Ada71938 Rice gene
C 23	38.2	3.7	5499	4	AAS46566	Aas46566 Tumour su
C 24	37.6	3.7	10732	3	AAA10594	Aaa10594 Gene enco
25	36.6	3.6	1948	4	AAK51957	Aak51957 Human pol
26	36.6	3.6	2207	11	ADM03188	Adm03188 Human cdb
27	36.6	3.6	3957	6	ABA98590	Abas98590 Human tes
28	36.6	3.6	3963	9	ADB80383	Adb80383 Human MDD
C 29	36.2	3.5	667	4	AAS41706	Aas41706 Genomic s
C 30	36.2	3.5	667	4	ABA06788	Abas06788 Human gen
C 31	36.2	3.5	667	6	ABV84125	Abv84125 Human pol
C 32	36.2	3.5	1443	4	AAS41704	Aas41704 Genomic s
C 33	36.2	3.5	1443	4	ABA06786	Abas06786 Human gen
C 34	36.2	3.5	1443	6	ABV84123	Abv84123 Human pol
C 35	35.8	3.5	4086	8	ACF72703	Acf72703 Staphyloc
C 36	35.8	3.5	4086	8	ACA20939	Aca20939 Prokaryot
C 37	35.8	3.5	4143	9	ADA32719	Ada32719 DNA encod
C 38	35.8	3.5	4488	2	AAV74539	Aav74539 Staphyloc
C 39	35.4	3.4	457	9	ACH21644	Ach21644 Human adu
C 40	35.2	3.4	3441	2	AAK40160	Aak40160 Gastric c
C 41	35.2	3.4	9402	4	AAS46671	Aas46671 Tumour su
C 42	35	3.4	1361	10	ADC03457	Adc03457 Wheat flo
C 43	35	3.4	1528	8	ACC46225	Acc46225 Human dit
C 44	35	3.4	1973	4	AAK52941	Aak52941 Human pol
C 45	35	3.4	7714	4	ABL20142	Ab120142 Drosophil

ALIGNMENTS

RESULT 1

AAF62446

ID AAF62446 standard; cdna; 1495 BP.

XX

AC AAF62446;

XX

DT 05-NOV-2001 (first entry)

XX

DE A thaliana VRN1 coding sequence.

XX

KW VRN1, vernalisation; flowering; crop; ss.

XX

OS Arabidopsis thaliana.

XX

FH Key Location/Qualifiers

FT CDS 269..1294

FT /tag= a

FT /product= "VRN1"

XX

PN WO200121822-A1.

XX

PD 29-MAR-2001.

XX

PF 13-SEP-2000; 2000WO-GB003525.

PR 17-SEP-1999; 99GB-00022071.

XX

PA (PLAN-) PLANT BIOSCIENCE LTD.

XX

PI Dean C, Levy YY;

XX

DR WPI; 2001-273467/28.

PT P-PSDB; AAB35491.

XX

PT Novel VRN1 polynucleotide sequence encoding a polypeptide which alters

PT vernalisation responses of plant in which VRN1 nucleic acid is expressed,

PT useful for influencing and assessing vernalization phenotype of plants.

XX

PS Claim 2; Fig 7; 91pp; English.

XX

CC The present invention provides the protein and coding sequences of

CC Arabidopsis thaliana VRN1. This protein is capable of altering the

CC vernalisation responses of a plant. Also provided are a number of PCR

CC primers used to isolate the sequences. The sequences are useful in the

Query Match	99.9%;	Score 1026;	DB 12;	Length 1026;
Best Local Similarity	100.0%;	Pred. No. 1e-311;		

Matches 1026; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGCCAGCGCCCTTTCTTCCATAAGTTGATTTTCTCATCACTATCCAGAAAAACGCTCG	60
Db	1	ATGCCAGCGCCCTTTCTTCCATAAGTTGATTTTCTCATCACTATCCAGAAAAACGCTCG	60
QY	61	AGGGTCCAGATGAAGTTTGTGAGTAAATTCAGGATGAGCTTTCGGTGTGCTGTGCACTC	120
Db	61	AGGGTCCAGATGAAGTTTGTGAGTAAATTCAGGATGAGCTTTCGGTGTGCTGTGCACTC	120
QY	121	ACAGTACCTCATGCTCATGTTTTCGCGTGTAGGACTAAGGAAGCTCAACAAAAATTTGG	180
Db	121	ACAGTACCTCATGCTCATGTTTTCGCGTGTAGGACTAAGGAAGCTCAACAAAAATTTGG	180
QY	181	TTTCAAGATGGTTGGCAAGATTTGTGTGACCGTTACTCCATTCGCAATGGTTATCTTTTG	240
Db	181	TTTCAAGATGGTTGGCAAGATTTGTGTGACCGTTACTCCATTCGCAATGGTTATCTTTTG	240
QY	241	ATTTTATAGATATGAAGAACTCTGCTTCAGCGTCTACATTTTCAATTTATCCCACTCT	300
Db	241	ATTTTATAGATATGAAGAACTCTGCTTCAGCGTCTACATTTTCAATTTATCCCACTCT	300
QY	301	GAGATCAATTTACCATTTCCACCGTCTCATGATTCGCTCACAAACCACTTCAACCGGCC	360
Db	301	GAGATCAATTTACCATTTCCACCGTCTCATGATTCGCTCACAAACCACTTCAACCGGCC	360
QY	361	CGTTTGTGTTGAAGACCTTGAAGATGAAGATGCGAGGTCACTTTCTTCTCTGTGTAC	420
Db	361	CGTTTGTGTTGAAGACCTTGAAGATGAAGATGCGAGGTCACTTTCTTCTCTGTGTAC	420
QY	421	CCATCAACCTTCTCTGAGTCTACAGTACCAAGCAAAAGGGTATGCTAGTTCAAGCATC	480
Db	421	CCATCAACCTTCTCTGAGTCTACAGTACCAAGCAAAAGGGTATGCTAGTTCAAGCATC	480
QY	481	CAAACTTCTTCTGAGTCTTAAAGTGAAGGCGCAACGCGCAACCCCAAAATACCT	540
Db	481	CAAACTTCTTCTGAGTCTTAAAGTGAAGGCGCAACGCGCAACCCCAAAATACCT	540
QY	541	AAAAAGAGGGAGGAGAAAGAAAATGCTGATCTGAGGAAATAAACTCATCAGTCCG	600
Db	541	AAAAAGAGGGAGGAGAAAGAAAATGCTGATCTGAGGAAATAAACTCATCAGTCCG	600
QY	601	CGAGATGATGATCCAGAGAACCTTCAAAGTTCTACAGAGTCTTCTGCGAGAAAGAGA	660
Db	601	CGAGATGATGATCCAGAGAACCTTCAAAGTTCTACAGAGTCTTCTGCGAGAAAGAGA	660
QY	661	ACGCTGACTGCAGAGAAAGAGAGAGGCCATCAATGCGAGCCAAACGTTCCGAACCA	720
Db	661	ACGCTGACTGCAGAGAAAGAGAGAGGCCATCAATGCGAGCCAAACGTTCCGAACCA	720
QY	721	AACCTTTCTTCTCAGAGTGGTTCTGCGACCACTCTATATACAGAGTTGCATCATGTAT	780
Db	721	AACCTTTCTTCTCAGAGTGGTTCTGCGACCACTCTATATACAGAGTTGCATCATGTAT	780
QY	781	CTTCTCTTCTGGTTCTGAGAAAGTACCTAAGTGGATCTCCGGTTCAATCAAGTCCAG	840
Db	781	CTTCTCTTCTGGTTCTGAGAAAGTACCTAAGTGGATCTCCGGTTCAATCAAGTCCAG	840
QY	841	CTTGGCGAGAAACAATGGCCCTGTTCCATGCTCTACAAAGCCGGAGAGCAAAATTCAGT	900
Db	841	CTTGGCGAGAAACAATGGCCCTGTTCCATGCTCTACAAAGCCGGAGAGCAAAATTCAGT	900
QY	901	CAAGGATGGTACGAATTTCACTCTAGAGAACCACTTAGGAGAGGAGACGCTCTGTGTTT	960
Db	901	CAAGGATGGTACGAATTTCACTCTAGAGAACCACTTAGGAGAGGAGACGCTCTGTGTTT	960
QY	961	GAGCTGCTCAGAACCCAGAGATTTTCTGTTTGAAGTGAAGCTTTCGAGTCAACGAGTAC	1020
Db	961	GAGCTGCTCAGAACCCAGAGATTTTCTGTTTGAAGTGAAGCTTTCGAGTCAACGAGTAC	1020
QY	1021	GTCTGA 1026	
Db	1021	GTCTGA 1026	

RESULT 3			
AAC39629			
ID	AAC39629	standard; DNA; 1510 BP.	
XX			
AC	AAC39629;		
XX			
DT	17-OCT-2000	(first entry)	
XX			
DE	Arabidopsis thaliana	DNA fragment SEQ ID NO: 25332.	
XX			
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-00301439.		
XX			
PR	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	30-APR-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	06-MAY-1999; 99US-0132487P.		
PR	07-MAY-1999; 99US-0132863P.		
PR	11-MAY-1999; 99US-0134256P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	18-MAY-1999; 99US-0134370P.		
PR	18-MAY-1999; 99US-0134768P.		
PR	19-MAY-1999; 99US-0134941P.		
PR	20-MAY-1999; 99US-0135124P.		
PR	21-MAY-1999; 99US-0135353P.		
PR	24-MAY-1999; 99US-0135629P.		
PR	25-MAY-1999; 99US-0136021P.		
PR	27-MAY-1999; 99US-0136392P.		
PR	28-MAY-1999; 99US-0136782P.		
PR	01-JUN-1999; 99US-0137222P.		
PR	03-JUN-1999; 99US-0137528P.		
PR	04-JUN-1999; 99US-0137502P.		
PR	07-JUN-1999; 99US-0137724P.		
PR	08-JUN-1999; 99US-0138094P.		
PR	10-JUN-1999; 99US-0138540P.		
PR	10-JUN-1999; 99US-0138847P.		
PR	14-JUN-1999; 99US-0139119P.		
PR	16-JUN-1999; 99US-0139452P.		
PR	16-JUN-1999; 99US-0139453P.		
PR	17-JUN-1999; 99US-0139492P.		
PR	18-JUN-1999; 99US-0139454P.		
PR	18-JUN-1999; 99US-0139455P.		
PR	18-JUN-1999; 99US-0139456P.		

99.8% score 1025.4: DB 3: Length 1510;

Db 406 ACAGTACCTGATGGTCATGTTGGCGGTGATAGGCTAAGGAAGCTGACCAAAAATTTGG 465
Qy 181 TTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTGCGATTGGTTATCTTTTG 240
Db 466 TTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTGCGATTGATTTATCTTTG 525
Qy 241 ATTTTATAGATGAAGGAACCTGCGCTTCAGGCTCTACATTTTCAATTTATCCCACTCT 300
Db 526 ATTTTATAGATGAAGGAACCTGCGCTTCAGGCTCTACATTTTCAATTTATCCCACTCT 585
Qy 301 GAGATCAATTACCATTTCCACCGTCTCATGGATTCGGCTCACAACCACTTCAAAACGCGC 360
Db 586 GAGATCAATTACCATTTCCACCGTCTCATGGATTCGGCTCACAACCACTTCAAAACGCGC 645
Qy 361 CGTTTCTTTGAAGACCTTGAAGATGAAGATGCGAGGTTCATCTTTCTCTCTCTGTGTAC 420
Db 646 CGTTTCTTTGAAGACCTTGAAGATGAAGATGCGAGGTTCATCTTTCTCTCTGTGTAC 705
Qy 421 CCATCAACCTTCTCTGAGTCTACAGTACGAGCCAAACAAAGGGTATGCTAGTTCAGCCATC 480
Db 706 CCATCAACCTTCTCTGAGTCTACAGTACGAGCCAAACAAAGGGTATGCTAGTTCAGCCATC 765
Qy 481 CAAACCTTCTCTGAGTCTACAGTACGAGCCAAACAAAGGGTATGCTAGTTCAGCCATC 540
Db 766 CAAACCTTCTCTGAGTCTACAGTACGAGCCAAACAAAGGGTATGCTAGTTCAGCCATC 825
Qy 541 AAAAAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 826 AAAAAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 885
Qy 601 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTGTCTTCGAGAAAGAGA 660
Db 886 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTGTCTTCGAGAAAGAGA 945
Qy 661 ACCGTGACTGCAG 720
Db 946 ACCGTGACTGCAG 1005
Qy 721 AACCTTCTCTCAGAGTGGTTCTGCGACCATCTATCTATACAGAGGTTGCAATCATGTAT 780
Db 1006 AACCTTCTCTCAGAGTGGTTCTGCGACCATCTATCTATACAGAGGTTGCAATCATGTAT 1065
Qy 781 CTTCTCTCTGGTTCCTGAGAGATACCTAAGTGGATCTCCGGGTTCATCAAAAGTCCAG 840
Db 1066 CTTCTCTCTGGTTCCTGAGAGATACCTAAGTGGATCTCCGGGTTCATCAAAAGTCCAG 1125
Qy 841 CTTGCGGAG 900
Db 1126 CTTGCGGAG 1185
Qy 901 CAAGGATGTTACGAATTCATCTAG 960
Db 1186 CAAGGATGTTACGAATTCATCTAG 1245
Qy 961 GAGCTGCTCAGAACCAAGAGATTTCTGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020
Db 1246 GAGCTGCTCAGAACCAAGAGATTTCTGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1305
Qy 1021 GTCTGAA 1027
Db 1306 GTCTGAA 1312

RESULT 4

AAC51591
ID AAC51591 standard; DNA; 1509 BP.

XX

AC AAC51591;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69076.

XX

Hybridisation assay; genetic mapping; gene expression control;
protein identification; signal transduction pathway; metabolic pathway;
promoter; termination sequence; ss.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-00301439.

25-FEB-1999; 99US-0121825P.

05-MAR-1999; 99US-0123180P.

09-MAR-1999; 99US-0123548P.

23-MAR-1999; 99US-0125788P.

25-MAR-1999; 99US-0126264P.

29-MAR-1999; 99US-0126785P.

01-APR-1999; 99US-0127462P.

06-APR-1999; 99US-0128234P.

08-APR-1999; 99US-0128714P.

16-APR-1999; 99US-0129845P.

19-APR-1999; 99US-0130077P.

21-APR-1999; 99US-0130449P.

23-APR-1999; 99US-0130891P.

28-APR-1999; 99US-0131449P.

30-APR-1999; 99US-0132048P.

04-MAY-1999; 99US-0132407P.

05-MAY-1999; 99US-0132484P.

06-MAY-1999; 99US-0132485P.

06-MAY-1999; 99US-0132486P.

07-MAY-1999; 99US-0132487P.

11-MAY-1999; 99US-0132863P.

14-MAY-1999; 99US-0134256P.

14-MAY-1999; 99US-0134218P.

14-MAY-1999; 99US-0134219P.

14-MAY-1999; 99US-0134221P.

18-MAY-1999; 99US-0134370P.

19-MAY-1999; 99US-0134768P.

20-MAY-1999; 99US-0134941P.

21-MAY-1999; 99US-0135124P.

21-MAY-1999; 99US-0135353P.

24-MAY-1999; 99US-0135629P.

25-MAY-1999; 99US-0136021P.

27-MAY-1999; 99US-0136392P.

28-MAY-1999; 99US-0136782P.

01-JUN-1999; 99US-0137222P.

03-JUN-1999; 99US-0137528P.

04-JUN-1999; 99US-0137502P.

07-JUN-1999; 99US-0137724P.

08-JUN-1999; 99US-0138094P.

10-JUN-1999; 99US-0138540P.

10-JUN-1999; 99US-0138847P.

14-JUN-1999; 99US-0139119P.

16-JUN-1999; 99US-0139452P.

16-JUN-1999; 99US-0139453P.

17-JUN-1999; 99US-0139492P.

18-JUN-1999; 99US-0139454P.

18-JUN-1999; 99US-0139455P.

18-JUN-1999; 99US-0139456P.

18-JUN-1999; 99US-0139457P.

18-JUN-1999; 99US-0139458P.

18-JUN-1999; 99US-0139460P.

18-JUN-1999; 99US-0139461P.

18-JUN-1999; 99US-0139462P.

18-JUN-1999; 99US-0139463P.

18-JUN-1999; 99US-0139750P.

21-JUN-1999; 99US-0139763P.

22-JUN-1999; 99US-0139817P.

23-JUN-1999; 99US-0139899P.

23-JUN-1999; 99US-0140353P.

PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.	Query Match	98.9%;	Score 1016;	DB 3;	Length 1509;
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.	Best Local Similarity	99.9%;	Pred. No. 1.8e-308;		
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.	Matches 1027;	Conservative	0;	Mismatches	0;
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.				Indels	1;
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.				Gaps	1;
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.					
PR	02-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.					
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.					
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.					
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.					
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.					
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.					
PR	13-JUL-1999;	99US-0143342P.	PR	04-OCT-1999;	99US-0157117P.					
PR	14-JUL-1999;	99US-0143624P.	PR	08-OCT-1999;	99US-0157753P.					
PR	15-JUL-1999;	99US-0144005P.	PR	05-OCT-1999;	99US-0157865P.					
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.					
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.					
PR	19-JUL-1999;	99US-0144352P.	PR	12-OCT-1999;	99US-0158369P.					
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.					
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.					
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.					
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.					
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.					
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.					
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.					
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.					
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.					
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.					
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.					
PR	22-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.					
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.					
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.					
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.					
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.					
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.					
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.					
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.					
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.					
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.					
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.					
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.					

Db 703 CCTAAAAGAGAGGGAGGAAGAAAATGCTGATCCTGAGGAATAAACTCATCAGCT 762
 Qy 598 CCGCAGATGATGATCCAGAGAACCGTTCAAAGTCTTACGAGAGTGTCTTCTGCGAGAAAG 657
 Db 763 CCGCAGATGATGATCCAGAGAACCGTTCAAAGTCTTACGAGAGTGTCTTCTGCGAGAAAG 822
 Qy 658 AGAACCGTGAATCCAGAGAAAGAGAGAGAGCCATCAATGACAGCCAAAACGTTTGAACCA 717
 Db 823 AGAACCGTGAATCCAGAGAAAGAGAGAGAGCCATCAATGACAGCCAAAACGTTTGAACCA 882
 Qy 718 ACAAAACCTTTCTTCTCAGAGTGGTTCTGCGAACCATCTATATACAGAGGTTGCATCATG 777
 Db 883 ACAAAACCTTTCTTCTCAGAGTGGTTCTGCGAACCATCTATATACAGAGGTTGCATCATG 942
 Qy 778 TATCTTCTCTGCGTTTCTGAGAGTACCTTAAGTGGATCTTCCGCGTTTCATCAAGTC 837
 Db 943 TATCTTCTCTGCGTTTCTGAGAGTACCTTAAGTGGATCTTCCGCGTTTCATCAAGTC 1002
 Qy 838 CAGCTTCCGAGAGAAACAAATGGCTGTTCGATGTCTTACAAGCCGGGAGAGCCAAATTC 897
 Db 1003 CAGCTTCCGAGAGAAACAAATGGCTGTTCGATGTCTTACAAGCCGGGAGAGCCAAATTC 1062
 Qy 898 AGTCAAGGATGGTACCAATTTCACTCTAGAGAACCACTTAGGAGAGGAGACGCTCTGTG 957
 Db 1063 AGTCAAGGATGGTACCAATTTCACTCTAGAGAACCACTTAGGAGAGGAGACGCTCTGTG 1122
 Qy 958 TTTGAGCTGCTCAGAACCCAGAGATTTCTGTTTGA 993
 Db 1123 TTTGAGCTGCTCAGAACCCAGAGATTTCTGTTTGA 1158

RESULT 6

AAF62416
 ID AAF62416 standard; DNA; 5000 BP.

XX AAF62416;

XX 05-NOV-2001 (first entry)
 XX A thaliana VRN1 gene.

XX VRN1; vernalisation; flowering; crop; ds.
 XX Arabidopsis thaliana.

XX WO200121822-A1.
 XX 29-MAR-2001.

XX 13-SEP-2000; 2000WO-GB003525.
 XX 17-SEP-1999; 99GB-00022071.

XX (PLAN-) PLANT BIOSCIENCE LTD.
 XX Dean C, Levy YY;

XX WP1; 2001-273467/28.

PT Novel VRN1 polynucleotide sequence encoding a polypeptide which alters
 PT vernalization response of plant in which VRN1 nucleic acid is expressed,
 PT useful for influencing and assessing vernalization phenotype of plants.

PS Claim 4; Page 70-73; 91pp; English.

CC The present invention provides the protein and coding sequences of
 CC Arabidopsis thaliana VRN1. This protein is capable of altering the
 CC vernalisation responses of a plant. Also provided are a number of PCR
 CC primers used to isolate the sequences. The sequences are useful in the
 CC production of crop plants, where they are able to control the timing of
 CC flowering, the duration of vernalisation required, the optimum
 CC temperature, or even eliminate the need for vernalisation completely. The
 CC present sequence is the VRN1 gene

XX SQ Sequence 5000 BP; 1447 A; 871 C; 931 G; 1750 T; 0 U; 1 Other;
 Query Match 63.6%; Score 653; DB 4; Length 5000;
 Best Local Similarity 77.1%; Pred. No. 5.7e-194;
 Matches 968; Conservative 1; Mismatches 1; Indels 285; Gaps 3;
 Qy 58 CTGAGGCTCCAGATAAAGTTTGTGAGTAAATTAAGAGATGAGCTTTCCGTTGCTGTGCA 117
 Db 3117 CAGAGGCTCCAGATAAAGTTTGTGAGTAAATTAAGAGATGAGCTTTCCGTTGCTGTGCA 3176
 Qy 118 CTCACAGTACCTGATGCTCATGTTGGCGTGTAGGACTAAGGAAGCTGACAAATAAT 177
 Db 3177 CTCACAGTACCTGATGCTCATGTTGGCGTGTAGGACTAAGGAAGCTGACAAATAAT 3236
 Qy 178 TCGTTTCAAGATGTTTGGCAAGATTTGTGACCGTTTACTCCATTCGCAATTCGTTATCTT 237
 Db 3237 TCGTTTCAAGATGTTTGGCAAGATTTGTGACCGTTTACTCCATTCGCAATTCGTTATCTT 3296
 Qy 238 TTGATTTTATGATATGAAGGAACTCTGCCTTCAGCGTCTACATTTTCAATTTATCCAC 297
 Db 3297 TTGATTTTATGATATGAAGGAACTCTGCCTTCAGCGTCTACATTTTCAATTTATCCAC 3356
 Qy 298 TCTGAGATCAATTAACATTTCCACCGTCTCATGATTCGGCTCAACAACCTTCAAACGC 357
 Db 3357 TCTGAGATCAATTAACATTTCCACCGTCTCATGATTCGGCTCAACAACCTTCAAACGC 3416
 Qy 358 GCCCGTTTGTGAAAGACCTTGAAGATGAAGATGCGAGGTCTATCTTCTCTCTCTG 417
 Db 3417 GCCCGTTTGTGAAAGACCTTGAAGATGAAGATGCGAGGTCTATCTTCTCTCTCTG 3476
 Qy 418 TACCCTCAACCTTCTGAGTCTACAGTACCAGCAACAAAGGATGATCTAGTTTCAGCC 477
 Db 3477 TACCCTCAACCTTCTGAGTCTACAGTACCAGCAACAAAGGATGATCTAGTTTCAGCC 3536
 Qy 478 ATCCAAACCTTGTTCAGTGGACCACTTAA----- 506
 Db 3537 ATCCAAACCTTGTTCAGTGGACCACTTAAAGGTGATATTTATAACCACTGATTCCTTT 3596
 Qy 507 ----- 506
 Db 3597 ATCTATCTGATTAAGCGTCTTATCTTCTTTCAGGTTGATGTTGATATTTCTTTA 3656
 Qy 507 -----AGCTGAAGAGCAACGCCAACCCCAAAATACCTTAAAGAGAGGAGGAAG 561
 Db 3657 TCTCCAGCTGAAGAGCAACGCCAACCCCAAAATACCTTAAAGAGAGGAGGAAG 3716
 Qy 562 AAAAATGCTGATCTT----- 576
 Db 3717 AAAAATGCTGATCTTCTGTAAGCACTTTTCTCTTTGAAATGCTTCAGACTCGTTTTT 3776
 Qy 577 -----GAGGA 582
 Db 3777 GGATTCACAGATTTCTCTCATGATATATATCTTTTGTATATTTGCTTACAGAGGA 3836
 Qy 583 ATAAACTCATCAGCTCCGCGAGATGATGATCCAGAGAACCGTTCAAAGTTCTTACGAGCT 642
 Db 3837 ATAAACTCATCAGCTCCGCGAGATGATGATCCAGAGAACCGTTCAAAGTTCTTACGAGCT 3896
 Qy 643 GCTTCTGCGAGAAAGAGAACCGTGTCTCAGAGAAAGAGAGAGAGCCATCAATCGAGCC 702
 Db 3897 GCTTCTGCGAGAAAGAGAACCGTGTCTCAGAGAAAGAGAGAGAGCCATCAATCGAGCC 3956
 Qy 703 AAAAGTTTGAACCAACAAACCTTTCTTTCAGAGTGGTTCTGCGACCATCTTATCTATAC 762
 Db 3957 AAAAGTTTGAACCAACAAACCTTTCTTTCAGAGTGGTTCTGCGACCATCTTATCTATAC 4016
 Qy 763 AGAGTTTGCATCAT----- 776
 Db 4017 AGAGTTTGCATCATGTTAATAAAAAACATCTTAGGAAGACTTAATCTTATCGGTGCTT 4076
 Qy 777 -----GTATCTTCTCTCTGG 792

Db 4077 CACTGATCTTAAAGAACGCTTCGTTTCTCTCTCAACAGTATCTTCTCTGGG 4136
Qy 793 TTTGCTGAGAGTACTTAAGTGGATCTCCGGGTTTCATCAAAGTCCAGCTTGGCGAAG 852
Db 4137 TTTGCTGAGAGTACTTAAGTGGATCTCCGGGTTTCATCAAAGTCCAGCTTGGCGAAG 4196
Qy 853 CAATGGCTGTTCGATGTCCTTACAAACCGGGAGAGCCAAATTCAGTCAAGGATGCTAC 912
Db 4197 CAATGGCTGTTCGATGTCCTTACAAACCGGGAGAGCCAAATTCAGTCAAGGATGCTAC 4256
Qy 913 GAAATTCAGTCTAGAGCAACCTTAGGAGAGAGAGAGCTCTGTGTGTTGAGCTGCTCAGA 972
Db 4257 GAAATTCAGTCTAGAGCAACCTTAGGAGAGAGAGAGCTCTGTGTGTTGAGCTGCTCAGA 4316
Qy 973 ACCAGAGATTTGTTTTGAAAGTGAACGCTTTTCGAGTCAACGAGTACGTCGTGAA 1027
Db 4317 ACCAGAGATTTGTTTTGAAAGTGAACGCTTTTCGAGTCAACGAGTACGTCGTGAA 4371

RESULT 7
AAC37903
ID AAC37903 standard; DNA; 753 BP.
XX AAC37903;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 19075.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128214P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132247P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.

```
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.

PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 53.4%; Score 548.4; DB 3; Length 753;
Best Local Similarity 99.8%; Pred. No. 1.7e-161;
Matches 549; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 478 ATCCAAACCTTGTTCACCTGGACCAAGTAAAGCTGAAGCCCAACGCGCAACCCCAAAAATA 537
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 AACCAACCTTGTTCACCTGGACCAAGTAAAGCTGAAGCCCAACGCGCAACCCCAAAAATA 60

QY 538 CCTAAAAGAGAGCGGAGGAAGAAAATGCTGATCTCTGAGGAAATAAATCATCAGCT 597
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CCTAAAAGAGAGCGGAGGAAGAAAATGCTGATCTCTGAGGAAATAAATCATCAGCT 120

QY 598 CGCGGAGATGATCCAGAGAACCGTTCAAAAGTTCTACGAGAGTGTCTCTCGGAGAAAG 657
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 CGCGGAGATGATGATCCAGAGAACCGTTCAAAAGTTCTACGAGAGTGTCTCTCGGAGAAAG 180

QY 658 AGAACCGTGACTGCAGAGAAAGAGAGAGAGCCATCAATGCAGCCCAAAACGTTCCGAACCA 717
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 AGAACCGTGACTGCAGAGAAAGAGAGAGAGCCATCAATGCAGCCCAAAACGTTCCGAACCA 240

QY 718 ACAAAACCTTCTTCAGAGTGTCTTCGGAACAATCCTATCTATACAGAGTTGCATCATG 777
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 ACAAAACCTTCTTCAGAGTGTCTTCGGAACAATCCTATCTATACAGAGTTGCATCATG 300

QY 778 TATCTTCCTCTCGGTTTCTGAGAGTACCTAAGTGGATCTCCGGGTTTCATCAAGTC 837
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 TATCTTCCTCTCGGTTTCTGAGAGTACCTAAGTGGATCTCCGGGTTTCATCAAGTC 360

QY 838 CAGCTTGGCGAGAAAACAATGGCTCTTCGATGCTCTACAAAGCCGGAGAGCCAAATTC 897
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 CAGCTTGGCGAGAAAACAATGGCTCTTCGATGCTCTACAAAGCCGGAGAGCCAAATTC 420

QY 898 AGTCAAGATGTTACGAATTCCTCTAGAGAACAACTTAGGAGAGAGAGACGCTCTGTGTG 957
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 AGTCAAGATGTTACGAATTCCTCTAGAGAACAACTTAGGAGAGAGAGACGCTCTGTGTG 480

QY 958 TTTGAGCTGCTCAGAACAGAGATTCGTTTGAAGTGACAGCCTTTTCGACTCAACGAG 1017
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 TTTGAGCTGCTCAGAACAGAGATTCGTTTGAAGTGACAGCCTTTTCGACTCAACGAG 540

QY 1018 TACGCTCGAA 1027
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 TACGCTCGAA 550

RESULT 8
AAC54934
ID AAC54934 standard; DNA; 1394 BP.
XX AAC54934;
AC AAC54934;
XX
XX
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 79593.
DE Arabidopsis thaliana
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
```

PR	23-MAR-1999;	99US-0125788P.	PR	16-JUL-1999;	99US-0144085P.
PR	25-MAR-1999;	99US-0126264P.	PR	16-JUL-1999;	99US-0144086P.
PR	29-MAR-1999;	99US-0126785P.	PR	19-JUL-1999;	99US-0144325P.
PR	01-APR-1999;	99US-0127462P.	PR	19-JUL-1999;	99US-0144331P.
PR	06-APR-1999;	99US-0128234P.	PR	19-JUL-1999;	99US-0144332P.
PR	08-APR-1999;	99US-01288714P.	PR	19-JUL-1999;	99US-0144333P.
PR	16-APR-1999;	99US-0129854P.	PR	19-JUL-1999;	99US-0144334P.
PR	19-APR-1999;	99US-0130077P.	PR	19-JUL-1999;	99US-0144335P.
PR	21-APR-1999;	99US-0130449P.	PR	20-JUL-1999;	99US-0144352P.
PR	23-APR-1999;	99US-0130510P.	PR	20-JUL-1999;	99US-0144632P.
PR	28-APR-1999;	99US-0130891P.	PR	20-JUL-1999;	99US-0144884P.
PR	30-APR-1999;	99US-0131449P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0144806P.
PR	04-MAY-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	05-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145085P.
PR	06-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145089P.
PR	07-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145192P.
PR	11-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145145P.
PR	11-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134219P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	99US-01351124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	17-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139750P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139763P.	PR	22-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139817P.	PR	23-AUG-1999;	99US-0149930P.
PR	21-JUN-1999;	99			

PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161933P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 42.9%; Score 441; DB 3; Length 1394;			
Best Local Similarity 73.5%; Pred. No. 1.5e-127;			
Matches 646; Conservative 0; Mismatches 195; Indels 38; Gaps 5;			
QY	151	GGACTAAGAAAGCTGACAAACAAATTTGGTTTCAAGATGGTTGGCAAGAGTTTGTGAC	210
DB	278	GGACTAAGAAAGCCAAACAAATCTGTTTCAAGACGGTTGGCAGGAGTTGTCAAC	337
QY	211	CGTTACTCCATTCGATTCGGTTATCTTTTGGATTTTATGATATGAAGAACTCTGCCCTC	270
DB	338	CGTTTCTCCATTCGATTCGGTTATCTTTTGGATTTTATGATATGAAGAACTCTGCCCTC	393
QY	271	AGCGTCTACATTTTCAATTTATCCACTCTGAGATCAATTTACCATTCACCGGTCTCATG	330
DB	394	ATCTCTCCACTCTGATGATCAACACC-----ATTCTAGTAGTGAAGCTTTATGCAATG	449
QY	331	GATTCGGCTCAACACCATT---CAAACCGCCCGTTTGGTTTGAAGACCTTCGAAGATGAA	387
DB	450	GATTCGGCACAGATCAGTTTCAACAAACGTGCTCGATTTGTTGAAGATCCTGAACTCAA	509
QY	388	GATCGGAGGTATCTTTCTCTCTGTTGTTACCCATCACCACCTTCTGTGAGTCTACGTA	447
DB	510	GATGCTAAGGTCAATTTATC-----CATCGAACCTTGAATCTACTGAA	551
QY	448	CCAGCCCAACAAAGGGTATCTAGTTTCAGCCATCCAAACCTTGTCTACTGGACCAGTTAA	507
DB	552	CCAGTGAATTAAGTTATGCGCGTTCTACAGCATCCAAAGCTTTTCAAAGATCTAAA	611
QY	508	GCTGAAGAGCCCAACGCCAACCCCAAAATACCTTAAAGAGAGGGAGGAAGAAAGAAAT	567
DB	612	GCTGAAGA-----AAGCCCAAGGTACTTTAAGNAGAGAGGAAGGAAGAAGAAAT	662
QY	568	GCTGATCTGAGGAATAAATCACTCAGCTCGCGAGATGATGATCCAGAGAACCGTTCA	627
DB	663	CCTAATCCCGAGGAGTAATCACTTTCACTCCCGGTGGAGATGACTCAGAGAACCGCTCA	722
QY	628	AAGTTCTACGAGAGTCTTCTGCGAGAAAGAGAACCGTCACTGCAAGAAAGAGAGAGA	687
DB	723	AAGTTCTACGAGAGTCTTCTGCTAGAAAGAGAACTGTAACCTGCAAGAAAGAGAGAGA	782
QY	688	GCCATCAATGCAGCCAAACGTTTCGAACCAACAAACCCCTTTTCTTCAGAGTGTGTTCTGCGA	747

Db	783	GCCTCAATGACGCCAAACATTCGAACCAACAAATCCTTACTTTTAGAGTTGTTCTGCGA	842
QY	748	CCATCTCTATATACAGAGTTGTCATCATGATCTTCTCTGGTTTCTGGAAGTAC	807
Db	843	CCATCATATCTATACAGAGTTGTCATCATGATCTTCTGGTTTCTGGAAGTAC	902
QY	808	CTAAGTGGGATCTCGGGTTTCATCAAGTCCAGTTCGGGAGAAACAATGGCCTGTTCGA	867
Db	903	CTAAGTGGGATCTCGGGTTTCATCAAGTCCAGTTCGGGAGAAACAATGGCCTGTTCGA	962
QY	868	TGCTCTTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGCTACGAATTCATCTAGAG	927
Db	963	TGCTCTTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGCTACGAATTCATCTAGAG	1022
QY	928	AACAACTTAGGAGAGGAGACCTCTGTGTGTTTGGCTGCTCAGAACCAAGATTTTCGTT	987
Db	1023	AACAATATAGGCAAGGAGATGATGTGTGTTTGGCTGCTCAGAACCAAGATTTTCGTT	1082
QY	988	TTGAAAGTGCACAGCCCTTTCGAGTCAACGAGTACGTCGA	1026
Db	1083	CTCGAAGTCACCGCCTTTCGTCATGATGTGTGA	1121
RESULT 9			
AAC53657			
ID	AAC53657 standard; DNA; 1396 BP.		
XX	AAC53657;		
XX	18-OCT-2000 (first entry)		
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 75262.		
DE	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
XX	EP1033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-00301439.		
XX	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	30-APR-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	07-MAY-1999; 99US-0132487P.		
PR	07-MAY-1999; 99US-0132863P.		
PR	11-MAY-1999; 99US-0134256P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	14-MAY-1999; 99US-0134370P.		

PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0137824P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 24-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140591P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143342P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 16-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155119P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-015753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.

PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161933P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match			
Best Local Similarity 73.5%; Score 441; DB 3; Length 1396;			
Matches 646; Conservative 0; Mismatches 195; Indels 38; Gaps 5;			
QY	151	GGACTAAGGAAGCTGACAAACAATAATTTGGTTTCAAGATGGTTGGCAAGAGTTTGTGAC	210
DB	279	GGACTAAGGAAGCCAAACAATAATCTGGTTTCAAGACGGTTGGCAGGAGTTTGTCAAC	338
QY	211	CGTTACTCCATTCGCATTCGTTATCTTTGATTTTAGATATGAAGAACTCGCTTC	270
DB	339	CGTTCTCCATTCGGATTCGT-----TTCAGATACAAAGTTACGTCTACATTTTCAATTT	394
QY	271	AGCGTCTACATTTTCAATTTATCCCACTCTGAGATCAATTTACCATTCACCGGTCTCATG	330
DB	395	ATCCTCCCACTCTGAGATCAACACC-----ATTCTAGTAGTGAAGCTCTTATGCAATG	450
QY	331	GATTCGGCTCACAAACACTT---CAAACGGCCGGTTTGTGGAAGACCTTGAAGATGAA	387
DB	451	GATTCGGCACAGAAATCAGTTTCAACAAACGTGCTCGATTTGTTGAAGATCCTGAACCTCAA	510
QY	388	GATGCCGGGTCTATCTTCTCTTCTGTGTACCCATCACCCTTCTGAGTCTACAGTA	447
DB	511	GATGCTAAGGTCAATTTATC-----CATCGAACCCCTGAATCTACTGAA	552
QY	448	CCAGCCAAACAAGGGTATGCTAGTTCAGCCATCCAAACCTTGTTCACCTGCACCACTTAA	507
DB	553	CCAGTCAATTAAGGTTATGGCGTTCTACAGCCATCCAAAGCTTTTCAAGATCTAAA	612
QY	508	GCTGAAGACCAACGCCAACCCCAAAAATACCTTAAAGAGAGGGAGGAAGAAATAAT	567
DB	613	GCTGAAGA-----AACGCCAAGGTACTTAAGAAGAGAGGAAGGAAGAAGAT	663
QY	568	GCTGATCCTGAGGAATTAACATCATGCTCCGCGAGATGATGATCCAGAGAACCGTTCA	627
DB	664	CCTAATCCCGAGGAAGTAACCTTCTCAATCCCGGTGGAGATGACTCAGAGAACCGCTCA	723
QY	628	AAAGTTCTACGAGAGTCTCTTCGCAAGAAAGAGAAACCGTGACTGCAGAAAGAGAGAGA	687
DB	724	AAAGTTCTACGAGAGTCTCTTCGCTAGAAAGAGAACTGTAACTGCAGAGAAAGAGAGA	783
QY	688	GCATCAATTCAGCCAAAACGGTTTCGAAACCAAAACCCCTTTCTTCAGAGTGGTTCTGCA	747
DB	784	GCCGTCAATTCAGCCAAAACATTCGAAACCAAAATCCTTACTTTAGAGTGTGTTCTGCA	843
QY	748	CCATCCTATCTATACAGAGTTGCATCATGTATCTTCTCTCTGGTTTCTGAGAGATAC	807
DB	844	CCATCATATCTATACAGAGTTGTCATCATGTATCTTGCCTATCTGGGTTTCTGAGAAATAC	903
QY	808	CTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGCTTTCGCGAGAAACAATGGCTGTCCA	867
DB	904	CTAAGTGGGATATCTGGTTTTCATCAAGTCCAGCTCGGTGAGAAACATGGCCAGTGAGG	963
QY	868	TGTCTCTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGGTACGAATTCATCTTAGAG	927
DB	964	TGCCTCTACAAAGCAGGAGAGCTAAGTTTTCAGCAAGGATGGTATGAGTTTCACTCCGAG	1023
QY	928	AACAACCTTAGGAGAGGAGAGCTGTGTGTTTGTAGCTCTCAGAACCAAGATTTTCGTT	987
DB	1024	AACAATATAGGGAAGGAGATGTATGTGTGTTTGTAGCTACTCAGAACTCGGATTTTCGTT	1083
QY	988	TTGAAAGTACAGCCCTTTTCGAGTCAACGAGTACGCTGA	1026
DB	1084	CTCAGAGTACCCGCCCTTTTCGTGTCAATGATGATGTGTA	1122

RESULT 10	
AAC37617	
ID	AAC37617 standard; DNA; 1344 BP.
XX	
AC	AAC37617;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 18041.
XX	
KW	Hybridization assay; genetic mapping; gene expression control;
KW	protein identification; signal transduction pathway; metabolic pathway;
KW	promoter; termination sequence; ss.
XX	
OS	Arabidopsis thaliana.
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	23-APR-1999; 99US-0130891P.
PR	30-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	04-MAY-1999; 99US-0132407P.
PR	05-MAY-1999; 99US-0132484P.
PR	06-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	07-MAY-1999; 99US-0132487P.
PR	11-MAY-1999; 99US-0132863P.
PR	14-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.
PR	18-MAY-1999; 99US-0134370P.
PR	19-MAY-1999; 99US-0134768P.
PR	20-MAY-1999; 99US-0134941P.
PR	21-MAY-1999; 99US-0135124P.
PR	21-MAY-1999; 99US-0135353P.
PR	24-MAY-1999; 99US-0135629P.
PR	25-MAY-1999; 99US-0136021P.
PR	27-MAY-1999; 99US-0136392P.
PR	28-MAY-1999; 99US-0136782P.
PR	01-JUN-1999; 99US-0137222P.
PR	03-JUN-1999; 99US-0137528P.
PR	04-JUN-1999; 99US-0137502P.
PR	07-JUN-1999; 99US-0137724P.
PR	08-JUN-1999; 99US-0138094P.
PR	10-JUN-1999; 99US-0138540P.
PR	10-JUN-1999; 99US-0138847P.
PR	14-JUN-1999; 99US-0139119P.
PR	16-JUN-1999; 99US-0139452P.
PR	16-JUN-1999; 99US-0139453P.
PR	17-JUN-1999; 99US-0139492P.
PR	18-JUN-1999; 99US-0139454P.
PR	18-JUN-1999; 99US-0139455P.
PR	18-JUN-1999; 99US-0139456P.
PR	18-JUN-1999; 99US-0139457P.
PR	18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P;
 PR 18-JUN-1999; 99US-0139460P;
 PR 18-JUN-1999; 99US-0139461P;
 PR 18-JUN-1999; 99US-0139462P;
 PR 18-JUN-1999; 99US-0139463P;
 PR 18-JUN-1999; 99US-0139750P;
 PR 18-JUN-1999; 99US-0139763P;
 PR 21-JUN-1999; 99US-0139817P;
 PR 22-JUN-1999; 99US-0139899P;
 PR 23-JUN-1999; 99US-0140353P;
 PR 23-JUN-1999; 99US-0140354P;
 PR 24-JUN-1999; 99US-0140695P;
 PR 28-JUN-1999; 99US-0140823P;
 PR 29-JUN-1999; 99US-0140911P;
 PR 30-JUN-1999; 99US-0141287P;
 PR 01-JUL-1999; 99US-0141842P;
 PR 01-JUL-1999; 99US-0142154P;
 PR 02-JUL-1999; 99US-0142055P;
 PR 06-JUL-1999; 99US-0142390P;
 PR 08-JUL-1999; 99US-0142803P;
 PR 09-JUL-1999; 99US-0142920P;
 PR 12-JUL-1999; 99US-0142977P;
 PR 13-JUL-1999; 99US-0143542P;
 PR 14-JUL-1999; 99US-0143624P;
 PR 15-JUL-1999; 99US-0144005P;
 PR 16-JUL-1999; 99US-0144085P;
 PR 16-JUL-1999; 99US-0144086P;
 PR 19-JUL-1999; 99US-0144325P;
 PR 19-JUL-1999; 99US-0144331P;
 PR 19-JUL-1999; 99US-0144332P;
 PR 19-JUL-1999; 99US-0144333P;
 PR 19-JUL-1999; 99US-0144334P;
 PR 19-JUL-1999; 99US-0144335P;
 PR 20-JUL-1999; 99US-0144352P;
 PR 20-JUL-1999; 99US-0144632P;
 PR 21-JUL-1999; 99US-0144884P;
 PR 21-JUL-1999; 99US-0144814P;
 PR 21-JUL-1999; 99US-0145086P;
 PR 21-JUL-1999; 99US-0145088P;
 PR 22-JUL-1999; 99US-0145085P;
 PR 22-JUL-1999; 99US-0145087P;
 PR 22-JUL-1999; 99US-0145089P;
 PR 22-JUL-1999; 99US-0145192P;
 PR 23-JUL-1999; 99US-0145145P;
 PR 23-JUL-1999; 99US-0145218P;
 PR 23-JUL-1999; 99US-0145219P;
 PR 26-JUL-1999; 99US-0145224P;
 PR 26-JUL-1999; 99US-0145276P;
 PR 27-JUL-1999; 99US-0145913P;
 PR 27-JUL-1999; 99US-0145918P;
 PR 27-JUL-1999; 99US-0145919P;
 PR 28-JUL-1999; 99US-0145951P;
 PR 02-AUG-1999; 99US-0146386P;
 PR 02-AUG-1999; 99US-0146388P;
 PR 02-AUG-1999; 99US-0146389P;
 PR 03-AUG-1999; 99US-0147038P;
 PR 04-AUG-1999; 99US-0147204P;
 PR 04-AUG-1999; 99US-0147202P;
 PR 05-AUG-1999; 99US-0147192P;
 PR 05-AUG-1999; 99US-0147260P;
 PR 06-AUG-1999; 99US-0147303P;
 PR 06-AUG-1999; 99US-0147416P;
 PR 09-AUG-1999; 99US-0147493P;
 PR 09-AUG-1999; 99US-0147935P;
 PR 10-AUG-1999; 99US-0148171P;
 PR 11-AUG-1999; 99US-0148319P;
 PR 12-AUG-1999; 99US-0148341P;
 PR 13-AUG-1999; 99US-0148565P;
 PR 13-AUG-1999; 99US-0148684P;
 PR 16-AUG-1999; 99US-0149368P;
 PR 17-AUG-1999; 99US-0149175P;
 PR 18-AUG-1999; 99US-0149426P;
 PR 20-AUG-1999; 99US-0149722P;
 PR 20-AUG-1999; 99US-0149723P;

PR 20-AUG-1999; 99US-0149929P;
 PR 23-AUG-1999; 99US-0149902P;
 PR 23-AUG-1999; 99US-0149930P;
 PR 25-AUG-1999; 99US-0150586P;
 PR 26-AUG-1999; 99US-0150884P;
 PR 27-AUG-1999; 99US-0151065P;
 PR 27-AUG-1999; 99US-0151066P;
 PR 27-AUG-1999; 99US-0151080P;
 PR 30-AUG-1999; 99US-0151303P;
 PR 31-AUG-1999; 99US-0151438P;
 PR 01-SEP-1999; 99US-0151930P;
 PR 07-SEP-1999; 99US-0152363P;
 PR 10-SEP-1999; 99US-0153070P;
 PR 13-SEP-1999; 99US-0153758P;
 PR 15-SEP-1999; 99US-0154018P;
 PR 16-SEP-1999; 99US-0154039P;
 PR 20-SEP-1999; 99US-0154779P;
 PR 22-SEP-1999; 99US-0155139P;
 PR 23-SEP-1999; 99US-0155486P;
 PR 24-SEP-1999; 99US-0155659P;
 PR 28-SEP-1999; 99US-0156458P;
 PR 29-SEP-1999; 99US-0156596P;
 PR 04-OCT-1999; 99US-0157117P;
 PR 05-OCT-1999; 99US-0157753P;
 PR 06-OCT-1999; 99US-0157865P;
 PR 07-OCT-1999; 99US-0158029P;
 PR 08-OCT-1999; 99US-0158232P;
 PR 12-OCT-1999; 99US-0158389P;
 PR 13-OCT-1999; 99US-0159293P;
 PR 13-OCT-1999; 99US-0159294P;
 PR 14-OCT-1999; 99US-0159295P;
 PR 14-OCT-1999; 99US-0159329P;
 PR 14-OCT-1999; 99US-0159330P;
 PR 14-OCT-1999; 99US-0159331P;
 PR 14-OCT-1999; 99US-0159637P;
 PR 14-OCT-1999; 99US-0159638P;
 PR 18-OCT-1999; 99US-0159584P;
 PR 21-OCT-1999; 99US-0160741P;
 PR 21-OCT-1999; 99US-0160767P;
 PR 21-OCT-1999; 99US-0160768P;
 PR 21-OCT-1999; 99US-0160770P;
 PR 21-OCT-1999; 99US-0160814P;
 PR 21-OCT-1999; 99US-0160815P;
 PR 22-OCT-1999; 99US-0160980P;
 PR 22-OCT-1999; 99US-0160981P;
 PR 22-OCT-1999; 99US-0160989P;
 PR 25-OCT-1999; 99US-0161404P;
 PR 25-OCT-1999; 99US-0161405P;
 PR 25-OCT-1999; 99US-0161406P;
 PR 26-OCT-1999; 99US-0161359P;
 PR 26-OCT-1999; 99US-0161360P;
 PR 26-OCT-1999; 99US-0161361P;
 PR 28-OCT-1999; 99US-0161920P;
 PR 28-OCT-1999; 99US-0161992P;
 PR 28-OCT-1999; 99US-0161993P;
 PR 29-OCT-1999; 99US-0162142P;

Query Match 42.8%; Score 439.4; DB 3; Length 1344;
 Best Local Similarity 73.4%; Pred. No. 4.7e-127;
 Matches 645; Conservative 0; Mismatches 196; Indels 38; Gaps 5;
 QY 151 GGACTAAGGAAGCTGACACAAATTTGGTTTCAAGATGGTTGGCAAGAGTTTGTTCAC 210
 |||||
 Db 226 GGACTAAGGAAGCCACAAATCTGGTTTCAAGACGGTTGGCAGGAGTTTGTCAAC 285
 |||||
 QY 211 CGTTACTCCATTCCGATTTGGTTATCTTTTGAATTTTATAGATGAGGAACCTGCTCCTC 270
 |||||
 Db 286 CGTTTCTCCATTCCGATTTGGT----TTCAGATACAAAGTTACAGTCTACATTTTCAATTT 341
 |||||
 QY 271 AGCGTCTACATTTTCAATTTTATCCCACTCTGAGATCAATTACCATTCACCGGTCTCATG 330
 |||||
 Db 342 ATCTCCCACTCTGAGATCAACACC-----ATTCTAGTAGAGTCTTTATGCAATG 397
 |||||

331 GATTTCGCTCACAAACCACTT---CAAAACGCGCCGTTTGTGTAAGACCTTGAAGATGAA 387
398 GATTTCGCGCAGAAATCAGTTCAACAAACGTCGATTTGTTTGAAGATCTGAACTCAAA 457
388 GATGCGGAGGTGATCTTTCTCTTCTGTGTATCCCATCCACTTCTCGAGTCTACAGTA 447
458 GATGCTAAGGTCAATTATC-----CATCGAACCTGATTTTCTGAA 499
448 CAGCCCAACAAAGGATGATGTTAGTTTTCAGGCATCCAAACCTTGTTCAGTCCAGTTAAA 507
500 CCAAGTGAATAAAGGTTATGCGGTTCTACAGCCATCCAAAGCTTTTTCAAAGATCTAAA 559
508 GCTGAGAGCCACGACCAACCCCAAAATACCTTAAAGAGAGGAGGAGGAAGAAAT 567
560 GCTGAAAG-----AAGCGCCAAGGTACTTAAAGAGAGGAGGAAGAAAGAAAT 610
568 GCTGATCTGAGGAAATAAATCATCAGCTCCGAGATGATGATCATCAGAGAACGTTCA 627
611 CCTAATCCGAGGAAGTAATCTTCACTCCGCTGGAGATGACTCAGAGAACCGCTCA 670
628 AAGTTCTACGAGAGTCTTCTGCGAGAAAGAGAACCGTCACTGCGAGAAAGAGAGAGA 687
671 AAGTTCTACGAGAGTCTTCTGCTAGAAAGAGAACTGTAATGCGAGAGAAAGAGAGAGA 730
688 GCCATCAATGCGACCAAAAGCTTCGAAACCAACCAACCCCTTCTTCAAGTGGTTCTGCGA 747
731 GCGGTCATGTCAGCCCAAAACATTCGAAACCAACCAATCCTTACTTAGAGTTGTTCTGCGA 790
748 CCATCTCTATCTATACAGAGGTTGTCATCATGTATCTTCTCTGCGTTTCTGAGAGTAC 807
791 CCATCATATCTATACAGAGTTGTCATCATGTATCTTCTGCGTTTCTGAGAAATAC 850
808 CTAAGTGGGATCTCCGGTTTCATCAAAGTCCAGTTCGCGAGAAACAAATGGCTGTTCGA 867
851 CTAAGTGGGATATCTGTTTTCATCAAGCTCCAGCTCGGTGAGAGAAACAAATGGCCAGTGG 910
868 TGCTCTACAAAGCCGGGAGAGCAAAATTCAGTCAAGGATGTAAGAAATTCATCTAGAG 927
911 TGCTCTACAAAGCAGGAGAGCTAAGTTTATAGCAAGGATGTAAGTTCACCTCGAG 970
928 AACAACCTTAGGAGAGGAGAGCTGTGTGTGTTTGTAGCTGCTCAGAACCAAGATTTCTGTT 987
971 AACAATATAGGAGAGAGATGATGTGTGTGTTTGTAGCTACTCAGAACTCGGATTCGTT 1030
988 TTGAAGTGCACAGCTTTTCGAGTCAACGAGTACGTTCTGA 1026
1031 CTCGAAGTCAACGCTTTTCGTTGTCATGATGTGTGA 1069

RESULT 11

AB212468
ID AB212468 standard; DNA; 681 BP.
XX
AC AB212468;
XX

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 273.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026695.

XX 24-AUG-2000; 2000US-0227866P.

XX 26-JAN-2001; 2001US-0264647P.

XX 22-JUN-2001; 2001US-0300111P.

PA (SCRI) SCRIPPS RES INST.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Harper JP, Krops J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.

PS Claim 144; SEQ ID NO 273; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office

XX Sequence 681 BP; 217 A; 139 C; 162 G; 163 T; 0 U; 0 Other;

Query Match 38.7%; Score 397.6; DB 6; Length 681;

Best Local Similarity 76.6%; Pred. No. 4.6e-114; Mismatches 134; Gaps 3;

Matches 538; Conservative 0; Indels 30;

QY 328 ATGATTCGCTCACAACCACTT---CAAAACGCGCCGTTTGTGTAAGACCTTGAAGAT 384

Db 7 ATGATTCGCGACAGATCAGTTCAACAAACGTCGATTCGTTTGAAGATCTCGAATC 66

QY 385 GAAGTGGGAGTCACTTTCTTCTTCTGTGTATCCCATCACTTCTCGAGTCTACA 444

Db 67 AAAGATGTAGGTCAATTTATCATCGAAC-----CCTGAATCTACT 108

QY 445 GTACCAGCCAAACAAAGGTATGCTAGTTTCAGCCATCCAAACCTTCTCACTGGACAGTT 504

Db 109 GAACAGTGAATAAAGGTTATGGCGGTTCTACAGCCATCCAAAGCTTTTCAAGATCT 168

QY 505 AAAGTGAAGAGCCCAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 564

Db 169 AAAGTGAAGAG-----AACGCGCCAGGTACTTTAAGAGAGAGGAGGAAGAGAG 219

QY 565 AATGCTGATCTGAGGAAATAAATCATCATGCTCCGCGAGATGATATCCAGAGAACCT 624

Db 220 AATCCTAATCCGAGGAAATAAATCTTCAACTCCCGGTGGAGATGATCTCAGAGAACCC 279

QY 625 TCAAGTTCACGAGAGTCTTCTGCGAGAGAGAACCGTGACCTGCAGAGAGAGAGAG 684

Db 280 TCAAGTTCACGAGAGTCTTCTGCTAGAGAGAGAACTGTAACTGCAGAGAGAGAGAG 339

QY 685 AGAGCCATCAATGACCAACCAACGTTCCGAAACCAACCAACCAACCTTCTTCTGAGTGTCTG 744

Db 340 AGAGCCGTCATGACCAACCAACCAACCAACCAACCAACCAACCTTCTTCTGAGTGTCTG 399

QY 745 CGACCATCTATCTATACAGAGGTTGATCATGTATCTTCTTCTGAGTGTCTGAGAGAG 804

Db 400 CGACCATCTATCTATACAGAGGTTGATCATGTATCTTCTTCTGAGTGTCTGAGAGAG 459

QY 805 TACCTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGCTTGGGAGAGAGAGAGAGAGAG 864

Db 460 TACCTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGCTTGGGAGAGAGAGAGAGAGAG 519

QY 865 CGATGCTCTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 924

Db 520 AGGTGCTCTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 579

QY 925 GAGAACCACTTAGGAG 984

Db 580 GAGAACATATAGCGGAGGAGATGATGTGTGTTGAGTACTCAGAACTCGGGATTTC 639
QY 985 GTTTGAAAGTGACAGCCTTTTCAGTCAACGAGTAGCTCTGA 1026
Db 640 GTTCTCGAAGTCACCGCCTTTTCGTGTCATGAGTATGTGTGA 681

RESULT 12
ABX62122/c
ID ABX62122 standard; DNA; 426 BP.
XX
AC ABX62122;
XX
DT 25-FEB-2003 (first entry)
XX
DE Arabidopsis thaliana expressed sequence related polynucleotide #237.
XX
KW Transgenic plant; plant; genetically modified cell; environmental stress;
KW ribozyme creation; disease resistance; stress tolerance;
KW fungicide screening; insecticide screening; gene; ds.
XX
OS Unidentified.
XX
PN US2002040490-A1.
XX
PD 04-APR-2002.
XX
PF 26-JAN-2001; 2001US-00770423.
XX
PR 27-JAN-2000; 2000US-0178512P.
XX
PA (GORL/) GORLACH J.
PA (ANY/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUY/) YU Y.
PA (RAME/) RAMEKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHAW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX
DR WPI; 2003-110411/10.
XX
XX Novel Arabidopsis thaliana nucleic acid useful for identifying homologous
PT or related genes, and to create genetically modified and transgenic
PT organisms, such as plant cells and plants.
XX
PS Claim 1; SEQ ID NO 237; 43pp; English.
XX
CC The invention describes an Arabidopsis thaliana nucleic acid (I). The
CC polypeptide (II) encoded by (I), transgenic plant (III) or genetically
CC modified cell (IV) are useful for screening a candidate agent for its
CC biological effect, by combining the candidate agent with (II), (III) or
CC (IV), and determining the effect of the candidate agent on (II), (III) or
CC (IV). (I) is useful for identifying homologous or related genes, for
CC producing compositions that modulate the expression or function of its
CC encoded protein, for mapping functional regions of the protein, in
CC diagnosis, for studying associated physiological pathways, for genetic

CC manipulation of cells, preferably plant cells, in screening assays of
CC various plant strains to determine the strains that are capable of
CC withstanding a particular disease or environmental stress, for enhancing
CC or inhibiting production of biosynthetic product in a plant, for
CC producing polypeptides, as probes for the detection of mRNA in biological
CC samples, to generate additional copies of (I), to generate ribozymes or
CC oligonucleotides, as single stranded DNA probes or as triple-strand
CC forming oligonucleotides, and to create genetically modified and
CC transgenic organisms, such as plant cells and plants. (II) or (III) is
CC useful for introducing or improving disease resistance and stress
CC tolerance in plants, screening biological active agents, e.g.,
CC fungicides, insecticides, etc., and for elucidating biochemical pathways.
CC (III) is useful as crops for their enhanced disease resistance, enhanced
CC traits of interest, for screening programs, as crops which exhibit
CC enhanced tolerance to environmental stress, or to produce a factor. This
CC sequence represents a nucleic acid that may correspond to naturally
CC occurring Arabidopsis thaliana expressed sequences. Note: the sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?docID=999909770423
XX
SQ Sequence 426 BP; 127 A; 101 C; 82 G; 114 T; 0 U; 2 Other;
Query Match 28.8%; Score 295.4; DB 8; Length 426;
Best Local Similarity 99.0%; Pred. No. 5.2e-82;
Matches 296; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 729 CTTGAGAGTGGTCTCGGACCATCTATCTATACAGAGGTGTCATGTATCTTCCTTC 788
Db 426 CTTGAGAGTGGTCTCGGACCATCTATCTATACAGAGGTGTCATGTATCTTCCTTC 367
QY 789 TGGGTTTGTCTGAGAGTACCTTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGTTCGCGA 848
Db 366 TGGGTTTGTCTGAGAGTACCTTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGTTCGCGA 307
QY 849 GAAACATGGCTGTTCGATGTCTCTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATG 908
Db 306 GAAACATGGCTGTTCGATGTCTCTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATG 247
QY 909 GTACGAATTCATCTAGAGACAACTTAGGAGAGAGAGACCTCTGTGTGTTTGTAGCTGCT 968
Db 246 GTACGAATTCATCTAGAGACAACTTAGGAGAGAGAGACCTCTGTGTGTTTGTAGCTGCT 187
QY 969 CAGAACCCAGAGATTTGTTTTGAAAAGTGACAGCCCTTTTCAGTCAACGAGTACGTCTGAA 1027
Db 186 CAGAACCCAGAGATTTGTTTTGAAAAGTGACAGCCCTTTTCAGTCAACGAGTACGTCTGAA 128

RESULT 13
ABZ15168
ID ABZ15168 standard; DNA; 2000 BP.
XX
AC ABZ15168;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2973.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
XX
PR 26-JAN-2001; 2001US-0264647P.
XX
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.

xx Arabidopsis thaliana expressed sequence related polynucleotide #569.
 DE
 XX
 XX Transgenic plant; plant; genetically modified cell; environmental stress;
 KW ribozyme creation; disease resistance; stress tolerance;
 KW fumigicide screening; insecticide screening; gene; ds.
 KW

XX 27-JAN-2000; 2000US-0178512P.
XX (GORL/) GORLACH J.
XX (ANYV/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUYV/) YU Y.
PA (RAME/) RAMEAKA J G.
PA (PAGE/) PAGE A.
PA (WATH/) MATHW A V.
PA (LEDE/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX
XX WPI; 2003-110411/10.
XX
XX Novel Arabidopsis thaliana nucleic acid useful for identifying homologous
PT or related genes, and to create genetically modified and transgenic
PT organisms, such as plant cells and plants.
XX
PS Claim 1; SEQ ID NO 569; 43pp; English.
XX
XX The invention describes an Arabidopsis thaliana nucleic acid (I). The
CC polypeptide (II) encoded by (I), transgenic plant (III) or genetically
CC modified cell (IV) are useful for screening a candidate agent for its
CC biological effect, by combining the candidate agent with (II), (III) or
CC (IV), and determining the effect of the candidate agent on (II), (III) or
CC (IV). (I) is useful for identifying homologous or related genes, for
CC producing compositions that modulate the expression or function of its
CC encoded protein, for mapping functional regions of the protein, in
CC diagnosis, for studying associated physiological pathways, for genetic
CC manipulation of cells, preferably plant cells, in screening assays of
CC various plant strains to determine the strains that are capable of
CC withstanding a particular disease or environmental stress, for enhancing
CC or inhibiting production of biosynthetic product in a plant, for
CC producing polypeptides, as probes for the detection of mRNA in biological
CC samples, to generate additional copies of (I), to generate ribozymes or
CC oligonucleotides, as single stranded DNA probes or as triple-strand
CC forming oligonucleotides, and to create genetically modified and
CC transgenic organisms, such as plant cells and plants. (II) or (III) is
CC useful for introducing or improving disease resistance and stress
CC tolerance in plants, screening biological active agents, e.g.,
CC fungicides, insecticides, etc., and for elucidating biochemical pathways.
CC (III) is useful as crops for their enhanced diseased resistance, enhanced
CC traits of interest, for screening programs, as crops which exhibit
CC enhanced tolerance to environmental stress, or to produce a factor. This
CC sequence represents a nucleic acid that may correspond to naturally
CC occurring Arabidopsis thaliana expressed sequences. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=99909770423
XX
SQ Sequence 411 BP; 155 A; 92 C; 57 G; 107 T; 0 U; 0 Other;
Query Match 6.1%; Score 63; DB 8; Length 411;
Best Local Similarity 82.8%; Pred. No. 7e-09;
Matches 72; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 940 GAAGGAGAGCTGTGTGTGTTTGGCTGCTCAGAACCCAGAGATTTCGTTTGAAGTGACA 999

Db 411 GAAGGAGAGTGTATGTGTGTTTGGCTGCTCAGAACCTCGGGATTTCGTTCTCGAAGTCACC 352
QY 1000 GCCTTTTCGAGTCAACGAGTACGCTCTGA 1026
Db 351 GCCTTTTCGAGTCAATGAGTATGTGTGA 325
Search completed: December 30, 2004, 13:26:53
Job time : 605 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2004, 12:43:00 ; Search time 113 Seconds
(without alignments)
6460.002 Million cell updates/sec

Title: US-10-088-187A-10_COPY_269_1295
Perfect score: 1027
Sequence: ; 1 atgccacgcctttcttcca.....agtcacgagtagctgtgaa 1027

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2.6/prodata/1/ina/5A COMB.seq:
2: /cgn2.6/prodata/1/ina/5B COMB.seq:
3: /cgn2.6/prodata/1/ina/6A COMB.seq:
4: /cgn2.6/prodata/1/ina/6B COMB.seq:
5: /cgn2.6/prodata/1/ina/PCTUS COMB.seq:
6: /cgn2.6/prodata/1/ina/backfile1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	43.2	4.2	7218	1	US-08-232-463-14
C 2	43	4.2	832	4	US-09-621-976-2813
C 3	40.2	3.9	1141	4	US-09-806-708B-22
C 4	39.2	3.8	7218	1	US-08-232-463-14
C 5	35.8	3.5	4143	4	US-09-328-352-4006
C 6	35.8	3.5	4488	4	US-08-956-171B-228
C 7	35.8	3.5	4488	4	US-08-781-986A-228
C 8	35.2	3.4	3441	3	US-08-896-164-85
C 9	34.6	3.4	832	4	US-09-621-976-2813
C 10	34.6	3.4	1664976	4	US-08-916-421B-1
C 11	34.6	3.4	1664976	4	US-09-692-570-1
C 12	34.4	3.3	1494	4	US-09-107-532A-2145
C 13	34.2	3.3	710	4	US-09-270-767-29582
C 14	34.2	3.3	1421	4	US-09-270-767-13580
C 15	34.2	3.3	3492	3	US-08-923-992A-9
C 16	34.2	3.3	6908	4	US-09-620-312D-977
C 17	34	3.3	3025	4	US-08-961-527-168
C 18	34	3.3	18431	4	US-09-221-017B-1090
C 19	33.6	3.3	499	4	US-09-270-767-358
C 20	33.6	3.3	499	4	US-09-270-767-15640
C 21	33.6	3.3	2535	4	US-09-799-451-549
C 22	33.4	3.3	314	4	US-09-614-912-151
C 23	33.4	3.3	1734	4	US-09-248-796A-6833
C 24	33.2	3.2	152331	3	US-09-128-155-16
C 25	33.2	3.2	176373	3	US-09-128-155-17
C 26	32.8	3.2	418	4	US-09-270-767-11173
C 27	32.8	3.2	2061	4	US-09-107-532A-1348

28	32.6	3.2	2617	1	US-08-430-024-1	Sequence 1, Appl
29	32.6	3.2	2617	1	US-08-782-009-1	Sequence 1, Appl
30	32.6	3.2	2617	1	US-09-017-302-1	Sequence 1, Appl
31	32.6	3.2	3294	3	US-08-923-992A-7	Sequence 7, Appl
32	32.6	3.2	3384	3	US-08-923-992A-5	Sequence 5, Appl
33	32.6	3.2	4200	1	US-08-242-932-1	Sequence 1, Appl
34	32.6	3.2	4200	1	US-08-714-481-1	Sequence 1, Appl
35	32.6	3.2	4200	3	US-08-923-992A-1	Sequence 1, Appl
36	32.6	3.2	4200	5	PCT-US95-06111-1	Sequence 1, Appl
C 37	32.6	3.2	90050	3	US-09-245-041-5	Sequence 5, Appl
C 38	32.6	3.2	90050	4	US-09-358-058B-5	Sequence 5, Appl
C 39	32.6	3.2	90050	4	US-09-893-238-5	Sequence 5, Appl
40	32.4	3.2	485	3	US-09-177-325-4	Sequence 4, Appl
41	32.4	3.2	485	3	US-09-411-812A-4	Sequence 4, Appl
42	32.4	3.2	485	3	US-09-590-113-4	Sequence 4, Appl
C 43	32.4	3.2	1141	4	US-09-806-708B-22	Sequence 22, Appl
C 44	32.4	3.2	2922	2	US-08-683-262B-74	Sequence 74, Appl
C 45	32.4	3.2	2922	3	US-09-361-707-74	Sequence 74, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Iardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgt-Fls
US-08-232-463-14

Query Match 4.2%; Score 43.2; DB 1; Length 7218;


```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match      3.4%, Score 34.6; DB 4; Length 832;
Best Local Similarity 9.6%; Pred. NO. 0.99;
Matches 30; Conservative 153; Mismatches 127; Indels 1; Gaps 1;

QY    331 GATTCCGCTCACACCACTTCAAAGCGCCGGTTGTTCGAAGACCTTGGAAGATGAAGAT 390
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     18 KKKWSWYMWTKWYMKTYWRMRKRRKKKAWKYKWKTWTWYRYAMWGTYKKKAWCRTK 77
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    391 GCCGAGGTCTCTTCCTCTCTCTGTGTACCCTACCATCCTCTCTGCTACTGACGTACCA 450
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     78 TKKKKKGYMMWYMGWRSSVMAWWITWTGYAYRSMTWYRWCWKKKYAKYKTKTCYS 137
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    451 GCCAACAAAGGTATGCTAGTTCAGCCATCAAACCTTGTTCTACTGGACCGATTAAAGCT 510
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     138 SKGWTWKRKKAWTTWWKKKY-YWAATRYMWMCMWTKRWASWYCYMWMGKARKWSTW 196
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    511 GAAGAGCCAAGCCCAACCCAAAATACTTAAAAAGAGAGGGAGGAGAAGAAAAATGCT 570
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     197 RKGRSYASARSARCKCYSCSGAWSWKYMMRMWRGWATGAGWKAWRASCMVRRIYAGK 256
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    571 GATCCTGAGAAATAAATACTCATCAGCTCCGCGAGATGATCATCCAGAGAACCGCTCAAAG 630
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     257 SKTSYKSMWCVTRSWIKYC YTKARWTGYCYRKGGMWGKGRWYASKYMKWKKWVCWAR 316
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    631 TTCTACGAGAG 641
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     317 MYRSTGTTRAS 327
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
PATENT NO. 6503729
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916.421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)

```

Query Match 3.4%; Score 34.6; DB 4; Length 1664976;
Best Local Similarity 48.3%; Pred. No. 80;
Matches 97; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY	506	AAGCTGAGAGCCCAACGCCAACCCCAAAAAATACCTTAAAAAGAGAGGGGAGGAAGAAAA	565
Db	547953	AAGTTGGAGATTATCGGANTTAAGATAAATTTGGAGAGGCATACATTAACATTAA	547894
QY	566	ATGCTGATCCTCGAGGAATAAACTCATCAGCTCCGCGAGATGATCCAGAGAACCGTT	625
Db	547893	AAGATGGTAAATATAGAGAAATCTCATCAGGTTTTGGAAATGGTGTAGCTGTAGAGTCT	547834
QY	626	CAAAGTTCTACGAGAGTGTCTTCGGAGAAAGAGAACCGTGACTGCAGAGAAAGAGAGAGA	685
Db	547833	TATACAAAAATGGATGGGGGTTTGTATCATCAACATATAGTGTAGAGAGGAATCGAAA	547774
QY	686	GAGCCATCAATCAGCCAAA	706
Db	547773	AATCAATAATAAGCGTATA	547753
RESULT 11			
US-09-692-570-1/c			
; Sequence 1, Application US/09692570			
; Patent No. 6797466			
; GENERAL INFORMATION:			
; APPLICANT: Bult et al.			
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ			
; Patent No. 6797466			
; TITLE OF INVENTION: jannaschii			
; FILE REFERENCE: PB275C1			
; CURRENT APPLICATION NUMBER: US/09/692,570			
; CURRENT FILING DATE: 2003-01-14			
; PRIOR APPLICATION NUMBER: US 60/024,428			
; PRIOR FILING DATE: 1996-08-22			
; PRIOR APPLICATION NUMBER: US 08/916,421			
; PRIOR FILING DATE: 1997-08-22			
; NUMBER OF SEQ ID NOS: 20			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 1			
; LENGTH: 1664976			
; TYPE: DNA			
; ORGANISM: Methanococcus jannaschii			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (28222)..(28222)			
; OTHER INFORMATION: n equals a, t, c, or g			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (28257)..(28258)			
; OTHER INFORMATION: n equals a, t, c, or g			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (84773)..(84773)			
; OTHER INFORMATION: n equals a, t, c, or g			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (84808)..(84808)			
; OTHER INFORMATION: n equals a, t, c, or g			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (84812)..(84812)			
; OTHER INFORMATION: n equals a, t, c, or g			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (98120)..(98120)			
; OTHER INFORMATION: n equals a, t, c, or g			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (98159)..(98159)			
; OTHER INFORMATION: n equals a, t, c, or g			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (98239)..(98239)			
; OTHER INFORMATION: n equals a, t, c, or g			
; FEATURE:			
; NAME/KEY: misc feature			

LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (600992)..(600992)

OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g

US-09-107-532A-2145

Query Match 3.4%; Score 34.6; DB 4; Length 1664976;
 Best Local Similarity 48.3%; Pred. No. 80;
 Matches 97; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 506 AAGCTGAAGAGCGCAACGCCAACCCCAAAATACCTTAAAGAGAGGGAGGAGAGAAA 565
 DB 547953 AAGTTGGAGATTATCGGATATAAGAAATAAATTTTGGAGAGAGCAATACAAATATTA 547894

QY 566 ATGCTGATCTGAGGAAATAACTATCAGCTCCGCGAGATGATCCAGAGAACCGTT 625
 DB 547893 AAGATGGTAAATAGAGAAATCTCATCAGGTTTGGAAATGGTGTAGAGTCT 547834

QY 626 CAAAGTTCTACGAGAGTGCTTCTGCGAGAAAGAACCGTGACTCCAGAGAGAGAGA 685
 DB 547833 TATACAAAATGATGGGGGTTTGTATCAATCAATAGTGAAGAGGAAATCGAAA 547774

QY 686 GAGCCATCAATGAGCGCAAAA 706
 DB 547773 AACTCATATAAAGCGGTATA 547753

RESULT 12
 US-09-107-532A-2145/c
 ; Sequence 2145, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD/ROM ISO9660
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ariniello, Pamela Deneke
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 2145:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1494 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (B) LOCATION 1...1494
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2145:

US-09-107-532A-2145

Query Match 3.3%; Score 34.4; DB 4; Length 1494;
 Best Local Similarity 59.0%; Pred. No. 1.6;
 Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 573 TCCTGAGGAAATAAATCACTCAGCTCCGCGAGATGATGATCCAGAGAACCGTTCAAAGTT 632
 DB 513 TCTGATGAAATCGCAGAGAGAGCTCTCTCACTTGATCCGCCCGGTGTACATTCAGGTT 454

QY 633 CTACGAGAGTGCTTCTTCGAGAAAGAGAACCGTGACTGCA 672
 DB 453 CCAAGGGTTTCTTGCTGGGCCATAGATGACTGGGTCTGTA 414

RESULT 13
 US-09-270-767-29582
 ; Sequence 29582, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 29582
 ; LENGTH: 710
 ; TYPE: DNA
 ; ORGANISM: *Drosophila melanogaster*
 ; US-09-270-767-29582

Query Match 3.3%; Score 34.2; DB 4; Length 710;
 Best Local Similarity 45.9%; Pred. No. 1.2;
 Matches 117; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 389 ATGCCGAGGTCATCTTTCTTCTTCTGTGTACCATCACCCTCTCTGAGTCTACAGTAC 448
 DB 78 ATACAAAAACAGCCCAATCTCTGCTGTAAAAATTTGCCAATTTATAACAAAAATGATC 137

QY 449 CAGCCACAAAGGTTGATGTTTTCAGCCATCCAAACCTTTTTCACCTGGACAGTTAAAG 508
 DB 138 CAAATCAAAATAATATGCAAAACACTAAGTTGTTTAAAGGAGGAGCTGGAAATGAAGTTG 197

QY 509 CTGAAGAGCCAAACGCCAACCCCAAAATACCTAAAAAGAGAGGAGGAGGAGAAAAATG 568
 DB 198 CAAAACACACAGCCGAGCAAAATAAAAACTGCAATGGGCGGTAAAAACGAAAAAGCACTG 257

QY 569 CTGATCTCTGAGAAATAAATCACTCAGCTCCGCGAGATGATGATCCAGAGAACCGTTCAA 628
 DB 258 CTACCCCTAAGGTAGTAAATAATCCAAACACTCAAAATAATTTTCAAGATAGCAACAA 317

QY 629 AGTTCTACGAGAGTG 643
 DB 318 ACAITCAAAAGAAATG 332

RESULT 14
 US-09-270-767-13580
 ; Sequence 13580, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 13580
 ; LENGTH: 1421
 ; TYPE: DNA

RESULT 15
US-08-923-992A-9
; Sequence 9, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3492 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS

THIS PAGE BLANK (USPTO)


```

Db 322 ACCGTTCTGATGGTCAATGTTTGGCGGTGAGGAATAAAGAAAGTTGACAAACAGGCTTGG 381
Qy 181 TTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTCGCAATGGTTATCTTTTG 240
Db 382 TTTAAGGAGGGTTGGCAGGAATTTTGCAGAACCTTACTATATCCGTTGGTCTACTTTTG 441
Qy 241 ATTTTATAGATATGAAGAACTCTGCTTCACGGTCTACATTTTCAATTTTATCCACTCT 300
Db 442 ATTTTCAGATATGAAGGAATTTCCGTTTCAGTTTAGCATATTTGATTTTGTACAACTCT 501
Qy 301 GAGATCAATATACATTTCCAACCGTCTCATGGATTCGGTTCACCAACCACTTCAAAACGCC 360
Db 502 GAAATAAATATCAACAAATCTCTTGTGGTACTCAATACAATCTCGGAAGACAATAT 561
Qy 361 CGTTTGGTTTGAAGACCTTGAAGATGAAGATGCCGAGGTCACTTTCTCTTCTCTCTG 416
Db 562 CCATTTGAAGAACTTGAAGATGATGAATGTGTCTCTCCAGCGCTGCTCAATTTTGTGGT 621
Qy 417 -----GTACCCATCACCACCTTCCCTGAGT 439
Db 622 GGGTCAAACTTAATCTGATATAAATGGAGTGGAGAGTCAACCATCAGCACCCTAAGGT 681
Qy 440 CTACAGTACACCCCAACAAAGGTATGCTAGTTTCAGCC-----ATCCAAACCTTG 489
Db 682 GTTAATAATCAACCTATTTCGGGGTATGGCTTTTCACTTACTGTATTTGAATTTAAAGAT 741
Qy 490 TTCACCTGACGATTAAGCTGAAGACCAAGCCAAACCCCAAAATACCTTAAAGAGA 549
Db 742 TCTGTAATGAAGTGAATGGCAATTTCTTCAGGTGAGAGCTACCAAAATCTGAAAAACCA 801
Qy 550 GGGAGGAAGAAAGAAATGCTCATCTGAGGAAATAAATCACTACAGCTCCGCGAGATGAT 609
Db 802 GGGAGGAAGAAAGCAAGTTTGAGCTAGTGAAGAGGATTCATCTCTTGGACATGAGAT 861
Qy 610 GATCAGAGAACCGTTTCAAGTTTCAAGAGTGTCTTCGAGAGAAAGAACCGTGACT 669
Db 862 GACATGCAATGCGTAATAGATTTTATGAAAGTGTCTCCGCCAGGAAGAAATTTGTGACT 921
Qy 670 GCAGAGAAAGAGAGAGAGCCATCAATGCAGCCAAACGTTGCAACCAACCAACCCCTTC 729
Db 922 GCTGAAGAAAGAGAGAGAGCGATTAATGCAGCCAAAGCATTTTGAGCCTACTAACCCTTC 981
Qy 730 TTCAGAGTGTCTCGACCATCTCTATATACAGAGTTGATCATGTATCTCTCTCT 789
Db 982 TCAGAGTGTCTTCGGACCGTCTTATCTATACAGGGAGTGTATGTATGTTACTACCATCG 1041
Qy 790 GGGTTTGTGAGAGTACCTAAGTGGGATCTCCGGGTTTCATCAAAAGTCCAGCTTCGGA - 848
Db 1042 TGCTTTGCTGAGCAACATCTAAGCGGGTTTCAGGATTCATTAACCTTCAGCTCCGGAT 1101
Qy 849 --GAAACAAATGCGCTGTGATGTCTCTACAAAGCCGGAGAGCCAAATTCAGTCAAGA 906
Db 1102 GGTAGACAGCGCCCTGTTTCGATSCGGTTATAGAGGAGGAGGAGGCTAAGTTTCAGTCAAGA 1161
Qy 907 TGGTACGAATTCACCTCAGAGAACTTTAGAGAGGAGAGCGTCTGTGTGTGTGAGCTG 966
Db 1162 TGGTATGAATTTACATTTGGAGAACAAATTTGGGGAGAGGAGTGTCTGTGTCTTGAAGCTG 1221
Qy 967 CTCAGAACCAAGATTTTGGTTTGAAGTGAAGAGCTTTTCAGTCAACGA 1016
Db 1222 CTCAGATCGAGGGAATTTGTGCTCAAGTTCACCGTATTTTGGTGTATGGA 1271

```

RESULT 2
 US-09-938-842A-273
 ; Sequence 273, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE OF INVENTION: SAME, AND METHODS OF USE
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 273
 ; LENGTH: 681
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-938-842A-273

Query Match	38.7%	Score 397.6;	DB 9;	Length 681;
Best Local Similarity	76.6%	Pred. No. 1.2e-115;		
Matches	538;	Conservative	0;	Mismatches 134; Indels 30; Gaps 3;

```

Qy 328 ATGATTCGCTCACAACCACTT---CAAACGCGCCGCTTTGTTTGAAGACCTTGAAGAT 384
Db 7 ATGATTCGCGCACAGAATCAGTTCAACAAACGCTGCTCGATTTGTTTGAAGATCCTGAATC 66
Qy 385 GAAGATCGGAGGTCACTTCTTCTTCTGTGTACCCATCACCACCTTCTCTGAGTCTACA 444
Db 67 AAGATGCTAAGGTCAATTTATCCATCGAAC-----CCTGAATCTACT 108
Qy 445 GTACAGCCAAACAAAGGATGCTAGTTTCAGCCATCCAAACCTTTGTTTCACTTGGACCACTT 504
Db 109 GAACAGGTGAATAAAGGTTATGGCGTTTCTACAGCCATCCAAAGCTTTTCAAAGATCT 168
Qy 505 AAAGCTGAAGAGCCAAACCCAAACCCAAATACTTAAAGAGAGAGGAGGAGAGAGAGAAA 564
Db 169 AAAGCTGAAGA-----AACGCCAAGGTACTTAAAGAGAGAGAGGAGAGAGAGAG 219
Qy 565 AATGCTGATCTGAGGAATAAATCACTACAGCTCCGCGAGATGATATCCAGAGAACCGT 624
Db 220 AATCTTAATCCGAGGAGTAAATCTTCTTCACTCCGCTGGAGATGACTCAGAGAACCGC 279
Qy 625 TCAAAGTTCTACGAGAGTCTTCTCGAGAGAAAGAGAACCGTGACTGCGAGAGAGAGAGAG 684
Db 280 TCAAAGTTCTACGAGAGTCTTCTGCTAGAAAGAGAACTGTAACTGCGAGAGAGAGAGAG 339
Qy 685 AGAGCATCAATCAGAGCCAAACCGTTGAAACCAACAAACCCCTTTCTTCCAGAGTGGTTCTG 744
Db 340 AGAGCCGTCAATGCGAGCCAAACATTCGAAACCAACAAATCCTTACTTTTAGAGTTGTTCTG 399
Qy 745 CGACATCTCTATACAGAGTTCGATCATGTATCTTCCCTCTGCGGTTCGCTGAGAGAG 804
Db 400 CGACATCATATCTATACAGAGTTTCATCATGTATCTTGGCCATCTGGGTTTCTGAGAGAA 459
Qy 805 TACCTAAGTGGGATCTCCGGGTTTCATCAAAAGTTCAGAGTTTGGGAGAGAAACAATGGCCTGT 864
Db 460 TACCTAAGTGGGATCTCTGGTTTCATCAAGCTCCAGCTCGGTGAGAGAAACAATGGCAGTG 519
Qy 865 CGATGTCTTCAAAAGCGGAGAGCCAAATTCAGTCAAGATGATGATGATGATGATGATGATGAT 924
Db 520 AGGTGCTCTTCAAAAGCGGAGAGAGTAAAGTTTAGCCAAAGGATGATGATGATGATGATGATGAT 579
Qy 925 GAGAACACTTTCAGAGAGAGAGAGTCTGTCTGTGTTGAGCTGCTCAGAACCCAGAGATTTTC 984
Db 580 GAGAACAAATATAGCGAAGGAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 639
Qy 985 GTTTTGAAGTGAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAG 1026
Db 640 GTTCTCGAAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAG

```

RESULT 3
 US-09-938-842A-273

RESULT 5

US-10-767-795-1316
; Sequence 1316, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; NUMBER OF SEQ ID NOS: 2004-01-30
; SEQ ID NO 1316
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C1540_3
US-10-767-795-1316

Query Match 28.0%; Score 287.8; DB 17; Length 1111;
Best Local Similarity 61.3%; Pred. No. 1.9e-80;
Matches 573; Conservative 0; Mismatches 267; Indels 95; Gaps 3;

QY 1 ATGCCACGCCCTTCTTCATAGTTGATTTCTCATCCACTATCCAGAAAACGTCG 60
DB |||||
QY 269 ATGCCACGCCCTTCTTCATAGTTGATTTCTCATCCACTATCCAGAAAACGTCG 328
DB |||||
QY 61 AGGGTCCCAAGATTTGTGAGTAAATCAAGATGAGCTTTCGGTGTCTGTGCATC 120
DB |||||
QY 329 AGGATCCCGGATAACTTCGTAAAGAAATTCGGGACGAACTTCTGTGTCGCCGTCTC 388
DB |||||
QY 121 ACAGTACCTGATGCTCATGTTGGCGTGTAGACTAAGAAAGCTGACAAACAAATTTGG 180
DB |||||
QY 389 ACTGTTCTCTGACGGTCATGTTTGGCGTGTAGAAATTAAGAAAGCGCAACAAAGGTTGG 448
DB |||||
QY 181 TTTCAGATGTTGGCAAGATTTGTGACCGTTACTCCATTCGCATGCTGTTATCTTTTG 240
DB |||||
QY 449 TTTCAGAGGGTTGGCAGAGTTTCTAGAGCGGTACCATATCCGTTGTGGCTACGTACTG 508
DB |||||
QY 241 ATTTTATGATATGAAGAACTCTGCTTTCAGCGTCTACATTTTCAATTTATCCACTCT 300
DB |||||
QY 509 GTTTTCAGATATGAATGGAATCTTGTTCAGTAAACCGTATCCATTGA----- 559
QY 301 GAGATCAATTACCATTCACCGGTCTCATGGATTCGGTCAACAACCATTCAAACCGCGC 360
DB -----
QY 560 ----- 559
QY 361 CGTTTGTGTAAGACCTTGAAGATGAAGATCGGAGGTGCATCTTCTCTCTGTGTAC 420
DB |||||
QY 560 -----AGACTTGAAGATGAAGATCGATCATCTCTCCGGCACTGCAGAT 602
DB |||||
QY 421 CCATCACCACCTTCTCAGTCTACAGTACCAGCAACAAAGGGTGTAGTTTCAGCCATC 480
DB |||||
QY 603 TTGT-----TTTCTGCGTCTAAAGTTAATACTGATGAACTGGAGTGTGAATCAAT 656
DB |||||
QY 481 CAAACCTGTTTCACTGGACCAAGTTAAAGCTGAAGAGCCAAACGCCCAACCCCAAAATACCT 540
DB |||||
QY 657 TTTCTGCGCTTAAAGGGCATGAATAGCCAACTCTTTTCGAGGTGCGAGTGTACCAAAACA 716
DB |||||
QY 541 AAAAGAGAGGGGAGAGAAATGCTGATCTCTGAGGAATTAACCTCATCAGCTCG 600
DB |||||
QY 717 AACAGCTGGGAGAGAAAGCAGAGTTTGTATCATCGTGAACCGGATTCATCCGTTGA 776
DB |||||
QY 601 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGTGTCTTCTGCGAGAAAGAGA 660
DB |||||
QY 777 CGGGAATATGATGTGTATACAACTTTAGATTTATACGAAAGTCTTCAGCCGAAAGAGA 836
DB |||||
QY 661 ACCGTGACTGCAAGAAAGAGAGAGAGCCATCAATGACGCAAAACCGTTGCAACCAACA 720
DB |||||
QY 837 ACCGTGACAAACGAAAGAGAGAGGGCAATTAACGCGAGCCAAATCGTTTGAGCGGATG 896
DB |||||

QY 721 AACCTTTCTTCAGAGTGGTTCTGCGACCATCTCTATCTATACAGAGTTGCATCATGTAT 780
DB |||||
QY 897 AACCTTTCTTCAGAGTGGTTCTTACGACCATCATATCTATACCGGGATGTATTATGTAC 956
DB |||||
QY 781 CTTCTTCTGGTGGTTGCTGAGAGTACCTAAGTGGGATCTCCGGTTCATCAAGTCCAG 840
DB |||||
QY 957 TTACCATCATGCTTTCCTGAGAGTATCTAAGTGGGTTTCGGGATTTATTAAACTTCAG 1016
DB |||||
QY 841 CTTGCGGA---GAAACATGGCTTTCGATGCTCTCTACAAAGCCGGAGAGCCAAATTC 897
DB |||||
QY 1017 CTTCCAGATGGAGACAATGGCCGCTTCGATGTCGTATAGAGTGGCAAGCCCAAGTTT 1076
DB |||||
QY 898 AGTCAAGGATGCTACGAATTCACCTCTAGAGAACAA 932
DB |||||
QY 1077 AGTCAGGATGCTAGAGTTTACGTTGAGAAATA 1111
DB |||||

RESULT 6

US-10-767-795-1317
; Sequence 1317, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; NUMBER OF SEQ ID NOS: 2004-01-30
; SEQ ID NO 1317
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C1540_2
US-10-767-795-1317

Query Match 24.8%; Score 255.2; DB 17; Length 1062;
Best Local Similarity 73.0%; Pred. No. 4.6e-70;
Matches 356; Conservative 0; Mismatches 128; Indels 4; Gaps 2;

QY 543 AAAGAGAGGGAGAGAGAAATGCTGATCTCTGAGGAATAAATCAATCATCAGCTCCCG 602
DB |||||
QY 164 AAAAAGCGTGGGAGGAGCGGAAGTTTGTATCTTAACTGCGAGGATTCATCTGCTGAGC 223
DB |||||
QY 603 AGATGATGATCCAGAGAACCGTTTCAAAGTTCTTACGAGAGTCTTCTGCGAGAAAGAGAAC 662
DB |||||
QY 224 TGAAGATGATGCTGACATGCGCTTAGGTGCTATGAAAGTCTTCTGCGGAAAGAGAAC 283
DB |||||
QY 663 CGTGAATGCAAGAAAGAGAGAGAGCGCATCAATGCGAGCCAAAGCTTTCGAGAACCAACAA 722
DB |||||
QY 284 CGTGACAGCTGAAGAAAGAGAGAGAGCAATCAATGCGGCGCAAGCAATTTGAGCCCTACTAA 343
DB |||||
QY 723 CCCTTCTTCAGAGTGGTTCTGCGACCATCTCTATCTATACAGAGTTCATCATGTATCT 782
DB |||||
QY 344 CCCTTCTTCAGAGTGGTCTGCGACCATCATATCTGTACAGGGGATGCAATATGTACTT 403
DB |||||
QY 783 TCCTTCTGGGTTTCTGAGAAAGTACCTAAGTGGGATCTTCCGGTTT-CATCAAAGTCCAGC 841
DB |||||
QY 404 ACCGTGCTGCTTCTGAGAGAGCATCTAAGTGGGTTTCTGATTTCCATTAACCTTCAGC 463
DB |||||
QY 842 TTGCGGA---GAAACATGGCTGTTCGATGCTCTCTACAAAGCCGGAGAGAGCCAAATTC 898
DB |||||
QY 464 TTCTGATGGGAGACAGTGGTCTGTACGATCTCGTTATAAAGAGGCGCAAGCTTAAGTTCA 523
DB |||||
QY 899 GTCAAGATGCTACGATTTCACTCTAGAGAACACTTAGGAGAGAGAGACCTCTGTGTGT 958
DB |||||
QY 524 GTCAGGATGCTACGATTTTACATTTGAGAAATAAATTTGGGAGAGAGGATGTCTGTGTCT 583
DB |||||
QY 959 TTGAGCTGCTCAGAAACAGAGATTTTCTGTTTTTGAAGTGACAGCTTTTCGAGTCAACGAGT 1018
DB |||||

Db 584 TCAGCTGCTCAGATCGAGGAATCGTGCTCAAGTACCCTATTTCTGTGAAGGAAA 643
QY 1019 ACCTCTGA 1026
Db 644 GCCTGGA 651

RESULT 7

US-10-424-599-118105
; Sequence 118105, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 118105
; LENGTH: 2141
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_77658C.1
US-10-424-599-118105

Query Match 22.1%; Score 227.2; DB 16; Length 2141;
Best Local Similarity 71.4%; Pred. No. 6.2e-61;
Matches 314; Conservative 0; Mismatches 123; Indels 3; Gaps 1;
QY 590 CATCAGCTCGGAGATGATGATCCAGAGAACCGTTCAAGTTCACGAGAGTCTTCG 649
Db 1203 CCTCTCTCGGCACGAGAGGAGGTAGAAATGCGCTTTAGATTTTATGAAGTCTTCG 1262
QY 650 CGAGAAAGAGAACCGTGACTGCAGAGAAAGAGAGAGCCATCAATGCAGCCAAACCT 709
Db 1263 CAGAAAGAACTGTGACAGCTGAAGAAAGAGAGAGGGTCAACAGAAAGCAAGCAT 1322
QY 710 TCGAACCAACAAACCTTTCTCAGAGTGGTTCTGCGACCATCTCTATATACAGAGTT 769
Db 1323 TTGAACCATTAATCTTCTCTGCGAGTTGCTCGGCCCTCATATTTATAGGGAT 1382
QY 770 GCATATGATATCTTCTCTGCTGAGAGTACCTAAGTGGGATCTCCGGGTTCA 829
Db 1383 GCATATGATATCTGCCATCATGCTTTGCAGAGAGCATTTGAATGGAGTTTCAGGGTTTA 1442
QY 830 TCAAGTCCAGCT--TGGGAGAAACATGGCTGTTTCGATGCTCTACAAAGCCGGA 886
Db 1443 TTAACCTTCAGATCTCTAATGGTAGACGTGGCCGGTTCGCTGCCCTTTATTAAGAGGTA 1502

QY 887 GAGCCAAATTCAGTCAAGATGGTAGCAATTCACCTAGAGAACAACTTAGGAGAGGAG 946
Db 1503 GAGCCAAATTCAGTCAAGATGGTAGCAATTCACCTAGAGAACAACTTAGGAGAGG 1562
QY 947 AGCTCTGTGTGTGAGTCTGCTCAGAACCCAGAGATTCCTTTTGAAGTGCACAGCTTTC 1006
Db 1563 ATGCTGTGTGTGTGAGTCTGCTCAGAACCCAGAGATTCCTTTTGAAGTGCACAGCTTTC 1622
QY 1007 GAGTCAACGAGTCACTCA 1026
Db 1623 ATGTTACCGAGGATGGGA 1642

RESULT 8

US-10-424-599-97592
; Sequence 97592, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 97592
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_59139C.1
US-10-424-599-97592

Query Match 22.0%; Score 226.2; DB 16; Length 1760;
Best Local Similarity 72.6%; Pred. No. 1.1e-60;
Matches 307; Conservative 0; Mismatches 113; Indels 3; Gaps 1;
QY 602 GAGATGATGATCCAGAGAACCGTTCAAGTTCACGAGAGTCTTCGCGAGAGAGAA 661
Db 946 GTGAATAATGTGCCGATTAAGCCCTGGAGGTTTTATGAACTGCGCATCTGCAGGAAAAGTA 1005
QY 662 CGTGACTGCAGAGAAAGAGAGAGAGCCATCAATGCAGCCAAAACGTTTCCGAACCAACAA 721
Db 1006 CCGTGACAGCAGAGAGAGAGAGAGAGGGCAATTAATGCATCAAAAACATTTGAACCACTA 1065
QY 722 ACCCTTTCTCAGAGTGGTTCTGCGACCATCTCTATACAGAGTTGTCATCATGTATC 781
Db 1066 ATCTTTCTCGCGAGTGTCTCTGCGACCCCTCTATTTGTATAGGGATGCATAATGTATC 1125
QY 782 TTCTTCTCGGGTTTCTGAGAGTACCTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGC 841
Db 1126 TGCTTCTCGCTTCTGCTGAAAAGAAATTTGAATGGGGTTTCGGGATTCATCAAACTTCAGT 1185
QY 842 TTGC--GGAGAAACAAATGGCTGTTTCGATGCTCTACAAAGCCGGAGAGCCAAATTC 898
Db 1186 TGTCAACCGGTAGACAGTGGTCTGTTGCTGCTCTATAGGGAGGTGAGGCCAAGTTAA 1245
QY 899 GTCAAGGATGTGCAAAATTCACCTTAGAGAACAACTTAGGAGAGAGAGCGTCTGTGTGT 958
Db 1246 GCCAAGTGTGGTTTGAATTCACAGTGGAGAACAAATTTGGGAGAGGTGAGCTGTGTGT 1305
QY 959 TTGAGCTGCTCAGAACCGAGAGATTCGTTTGAAGTGCACGCTTTTCAGTCAAGCAT 1018
Db 1306 TTGAGCTCCTTAGAACGAGGAAGTTGTGCTGCAAGTTACGGTATTTTCGCGTAACCGAGG 1365
QY 1019 ACG 1021
Db 1366 ATG 1368

RESULT 9

US-10-021-323-9015
; Sequence 9015, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 9015

	;	LENGTH: 644	
	;	TYPE: DNA	
	;	ORGANISM: Gossypium hirsutum	
	;	FEATURE:	
	;	NAME/KEY: unsure	
	;	LOCATION: (1) -- (644)	
	;	OTHER INFORMATION: unsure at all n locations	
	;	OTHER INFORMATION: Clone ID: LIB3828-022-Q1-K6-B11	
	;	US-10-021-323-9015	
		Query Match	17.0%; Score 174.8; DB 17; Length 644;
		Best Local Similarity	57.1%; Pred. No. 1.6e-44;
		Matches 420; Conservative 0; Mismatches 223; Indels 92; Gaps 2	
Qy	65	TCCAGATAAGTTTGTGAGTAATATTCAAGAGTAGACTTTCGGTTGCTGTTGCACCTCACAG	124
Dd	2	TCCCGGATAACTTTTGTTAAGAAATTTGGGGACGAACCTTCTGTTGCTGCTGCTCCTCACTG	61
Qy	125	TACCTGATGTCATGTTTGGCGTGTAGGACTAAGGAAAGCTGACAACAAAAATTTGGTTTC	184
Dd	62	TTCTGTATGGTCATGTTTGGCGTGTAGGAATTAAGAAAAGCCGACAACAAGGTTTGGTTTC	121
Qy	185	AAGATGTTTGGCAAGAGTTTCTGCACCGTTACTCATTCGCATATGGTTATCTTTTGATTT	244
Dd	122	ACGAGGTTGGCAGGAATTTGTAGACGGTACCATATCCGTTTGGCTACTTACTGGTTT	181
Qy	245	TTAGATATGAAGGAAACTCTGCCCTTCACGCTACTCATTTTCAAATTTATCCCACTCTGAGA	304
Dd	182	TCAGATATGAATGGAATCTTGTGTTTCAGTAAACCGTATCCAATTGA-----	228
Qy	305	TCAATTACCATTCCACCGTCTCATGGATTCGGTCTCAACCACTTCAACGCGCCCCGTT	364
Dd	229	-----	228
Qy	365	TGTTTTGAAGACCTTTGAAGATGAAGATGCCGAGGTTCATCTTTCTCTTCTGTGTACCCAT	424
Dd	229	-----AGAACTTGAAGATGACGAATGCACTCTCTCGGCACCTGCAGAAATTTGT	275
Qy	425	CACCACCTTCCTGAGTCTACAGTACGAGCCAAAGGATGTAGTATGCTAGCTTACGCCATFCCAA	484
Dd	276	-----TTTCTGCGCTAAAGTTAATAACTGCATGAACGTGAGTGGTGAATCAATTTTC	329
Qy	485	CCTTGTTCCTGGACCAGTTTAAAGCTGAAGAGCCAAAGCCCAACCCCAAAAATACCTAAA	544
Dd	330	GTGCGTTTAAAGGCGATGAATAGCCATCTTTTCAGGTGCGAGTGTCTACCAAACCAACA	389
Qy	545	AGAGGGGAGGAAAGAAAAATGCTGATCTCGAGGAAATAAACTCATCAGCTCGCGGAG	604
Dd	390	AGCCTGGGAGGAAAGACAGAAATTTGATCAGCTGACAGATTCATCTGTGTGACACG	449
Qy	605	ATGATGATCCAGAGAACCGTTCAAAGTTCACGAGAGTGCCTTCGCGAGAAAGAGAACCG	664
Dd	450	AATATGATGTGTATACAACTTTTAGATTATATGAAGTGCTTCAGCCAGAAAGAGAACTG	509
Qy	665	TGACTGCGAGAGAAAGAGAGAGCCCATCAATGTCAGCCAAACGTTCCAAACCAACACC	724
Dd	510	TGACACGGAGAAAGAGAGAGAGCAATTAACGACGCCAAATCGTTTGTGAGCGGTGAACC	569
Qy	725	CTTTCTTTCAGAGTGGTTCTGCGACCATCCTATCTATACAGAGTTGCAATCATGTATCTTC	784
Dd	570	CTTTCTGAGAGTTGTCTTACGACCATCATATCTATACCGGGATGTTATGTACTTAC	629
Qy	785	CTTCTGGGTTTGCTG	799
Dd	630	CATCATGCTNTGCTG	644

RESULT 10
US-10-424-599-97591
; Sequence 97591, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J

```

; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OP INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 97591
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_59138C.1
US-10-424-599-97591

```

Query Match	15.3%	Score 157.2;	DB 16;	Length 628;
Best Local Similarity	65.3%;	Pred. No. 6.4e-39;		
Matches 231; Conservative	0;	Mismatches 123;	Indels 0;	Gaps 0;
Qy	489	GTTCACCTGGACCACTAAAGCTGAAGAGCCAAACGCCAACCCAAAAATACCTAAAGAGAG	548	
Db	248	GTCCACTGAAGAATTGAATTTGGGTTATTCTTAATTTCTAAACAGAGACGAGTTAATAAAC	307	
Qy	549	AGGGAGGAAAGAAAAAATGCTGATCCTGAGGAAATAAATCATATCAGCTCCGCGAGATGA	608	
Db	308	CGCAAAAAGACGGGAATCAGAACCCATATGCGAGGAACCTTCTGCTGATAAATGAAGA	367	
Qy	609	TGATCCAGAGAACCGTTCAAAGTTCTACAGAGGTGCTTCTGCGAGAAAGAGAACCGTGAC	668	
Db	368	GGAGGCAGAAATGCGCTATAGGTTTTATGAAAGTGCAATCTGCAAGGAAAACGAACCGTGAC	427	
Qy	669	TGCAGAAAGAAAGAGAGAGCCATCAATGCAGCCAAAACGGTTCGAAACCAACAAACCCCTTT	728	
Db	428	GGCAGAGAAAGAGAGAAAGGGCAATTAATGCATCAAAAACATTTTGAACCGACTAATCCCTTT	487	
Qy	729	CTTTCAGAGTGGTTCTGCGACCATCTATCTATACAGAGGTTGCAATCATGTATCTTCCCTTC	788	
Db	488	CTGCGAGTTGTCTCGGACCCCTCCTATTGTATAGGGATGCATATGATATCTGCGCTTC	547	
Qy	789	TGGGTTTTGCTGAGAAGTACCTAAAGTGGGATCTCCGGGTTTCATCAAGTCCAGCT	842	
Db	548	CACCTTTGCTGAAAAGAAATTTGAATGGGTTTTCGGGATTCATCAAACTTCAGCT	601	

```

RESULT 11
US-10-767-795-1320
; Sequence 1320, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kao, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Mole
; TITLE OF INVENTION: Plants and Uses
; FILE REFERENCE: 38-21(53534) B
; CURRENT APPLICATION NUMBER: US/10/767
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 1320
; LENGTH: 650
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3543
US-10-767-795-1320

```

Query Match 10.8%; Score 111; DB 17; Length 650;
Best Local Similarity 69.5%; Pred. No. 3.7e-24;
Matches 166, Conservative 0; Mismatches 70; Indels 3; Gaps 1

QY 151 GGACTAAGGAAAGGTGACACACAAAATTTGGTTCAAGATGGTTGGCAAGAGCTTTCTTGAC 210

Db 260 GGAATAAAGAAAGTTGACAAAGGTTTGGTTCAGGAAGGTTGGCAGGAGTTTCTAGAG 319
Qy 211 CGTTACTCCATTCCGATTTGTTTCTTTTTCATTTTATGATATGAAGGAAACTCTGCTTC 270
Db 320 CGTTACTACTATTCGTTGTTGCTAGTACTGTTTTCAGATACGAGGAAATTCGCTTC 379
Qy 271 AGCGTCTACATTTTCATTTTATCCACTCTGAGATCAATTTACCATTCCACCGTCTCATG 330
Db 380 AGTGTAGTATATTTAATTTGTATACTCGGAATAAACTATCAGACTAATGCCCTCGTT 439
Qy 331 GATTCCGCTCAGACCACTTCAACGCGCCGTTTGTTCAGACCTTTGAAGATGAAGA 389
Db 440 GGTACTCAATACATCAACGGAAC---AATATCCGTTTGAACAACTTGAAGATGATGA 495

RESULT 12

US-10-021-323-7926/c
; Sequence 7926, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 7926
; LENGTH: 602
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3828-022-Q1-N6-E11
US-10-021-323-7926

Query Match 9.9%; Score 101.4; DB 17; Length 602;
Best Local Similarity 73.7%; Pred. No. 4.1e-21;
Matches 129; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
Qy 847 GAGAAACAATGGCTCTGATGCTCTACAAAGCCGGAGAGCCAAATTCAGTCAAGCA 906
Db 578 GGGAGACAATGGCCCAATCGATGCTGTATAGAGTGGCAAGCCAAAGTTTAGTCAAGGA 519
Qy 907 TGTACGAATTCACCTCTAGAACAACTTAGGAGAAGGAGAGCTCTGTGTGTTGAGCTG 966
Db 518 TGTACGAGTTTACGTTGGAGATAATTTGGGTGAAGGAGATGCTGTATCTTCGAGCTG 459
Qy 967 CTCAGAACAGAGATTCGTTTGAAGTGACAGCTTTCGAGTCAACGAGTACG 1021
Db 458 CTCAGATCAAGGAATTCGCTTAAAGTTACCGTATTTTCGTTAAGGGAAGCG 404

RESULT 13

US-09-938-842A-2973
; Sequence 2973, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2973
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2973

Query Match 9.8%; Score 101; DB 9; Length 2000;
Best Local Similarity 88.3%; Pred. No. 1.2e-20;
Matches 121; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
Qy 96 TGAGCTTTCGGTTGCTGTTGCACCTACAGTACTGATGCTCATGTTTGGCGTGTAGACT 155
Db 1784 TGAGCTTTCAGTTGTTTTCACCTTACATATACCTGATGCTCATGTTT-GCATGTAGACT 1842
Qy 156 AAGGAAAGCTGACAAACAAAATTTGGTTTCAAGATGTTGGCAAGAGTTTGTTCACCGTTA 215
Db 1843 AAGGAAAGCCAAACAAAATCTGTTTTCAGACGTTTGGCAGGAGTTTGTCAACCGTTT 1902
Qy 216 CTCATTTCGATTGGTT 232
Db 1903 CTCATTTCGATTGGTT 1919

RESULT 14

US-09-938-842A-2973
; Sequence 2973, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2973
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2973

Query Match 9.8%; Score 101; DB 11; Length 2000;
Best Local Similarity 88.3%; Pred. No. 1.2e-20;
Matches 121; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
Qy 96 TGAGCTTTCGGTTGCTGTTGCACCTACAGTACTGATGCTCATGTTTGGCGTGTAGACT 155
Db 1784 TGAGCTTTCAGTTGTTTTCACCTTACATATACCTGATGCTCATGTTT-GCATGTAGACT 1842
Qy 156 AAGGAAAGCTGACAAACAAAATTTGGTTTCAAGATGTTGGCAAGAGTTTGTTCACCGTTA 215
Db 1843 AAGGAAAGCCAAACAAAATCTGTTTTCAGACGTTTGGCAGGAGTTTGTCAACCGTTT 1902
Qy 216 CTCATTTCGATTGGTT 232
Db 1903 CTCATTTCGATTGGTT 1919

```

RESULT 15
US-10-739-930-3743/c
; Sequence 3743, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 3743
; LENGTH: 2207
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURES:
; NAME/KEY: unsure
; LOCATION: (1)..(2207)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: GLYMA-23APR03-CLUSTER63260_1
US-10-739-930-3743

```

Query Match	8.7%	Score 89.6	DB 18	Length 2207
Best Local Similarity	55.9%	Pred. No. 5.3e-17		
Matches 170	Conservative 0	Mismatches 134	Indels 0	Gaps 0
Qy	15	CTTCCATAAGTCGATTTTCTCATCCACTCCTCAAGAAAAACGCTCAGGGTCCGAGATAA	74	
Db	1758	CTTTTTCAAGATCATTAACCGCTCACATGTTTCATGAAGGAAGCTAATGATCCCAATATA	1699	
Qy	75	GTTTGTGAGTAAATTCGAAGGATGAGCTTTCGGTTCGTGTGCACCTCACAGTACCTGATCG	134	
Db	1698	GTTTGTGAAGAAATATGGAATAACGCTACAAAATACTCTATTTCGAAGACTCCAAATCG	1639	
Qy	135	TCATGTTTGGCGGTAGGACTAAGGAAGCTGACAAATAATTTGGTTTCAAGATGGTTG	194	
Db	1638	TGCTGTAATGAAAATGATTTTGAAAAAACGTCACGGTAAAAATATGGTTTCAAAAGGGATG	1579	
Qy	195	GCAAGAGTTGTGTGACCGTTACTCCATTCGCATTTGGTTATCTTTTCATTTTATGATATCA	254	
Db	1578	GAAAGAGTTTGCAGAGTATCACTCTCTAGCTCATGGCCATCTTTTGGTTTTCAGATGGGA	1519	
Qy	255	AGGAAACTCTGCCTTCAGCGCTCTACATTTTCAATTTTATCCCACTCTGAGATCAAAATACCA	314	
Db	1518	TGTAACCTTCTCATTTTCAGGTACACATCTTTGATTTTGAGTGGCCTTAGAGATTGAGTACCC	1459	
Qy	315	TTCC	318	
Db	1458	TACC	1455	

Search completed: December 30, 2004, 17:28:09
Job time : 664 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2004, 09:53:25 ; Search time 3904 Seconds
(without alignments)
9585.964 Million cell updates/sec

Title: US-10-088-187A-10_COPY_269_1295

Perfect score: 1027
Sequence: 1 atgcacgcctttcttcca.....agtcacgagtagctgctgaa 1027

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1025.4	99.8	1307	3	CNS0A70B
2	1013.4	98.7	1441	3	EX822585 Arabidops
3	996.6	97.0	1375	3	EX823300 Arabidops
4	470.8	45.8	701	6	EX824523 Arabidops
5	441.4	42.9	1420	3	CD822418 EN25.045C
6	441.4	42.9	1490	3	EX816331 Arabidops
7	437.8	42.6	1346	3	EX815859 Arabidops
8	410.8	39.9	880	3	EX816536 Arabidops
9	401.8	39.1	589	1	EX81843 Arabidops
10	396.2	38.6	780	7	AV540912 Arabidops
11	387.7	37.7	591	1	CK120978 204014.p1
12	382.6	37.3	1041	5	AI996485 70166887
13	380.2	37.0	645	6	EX835682 EX835682
14	362.6	35.3	906	7	CD823066 EN25.047I
15	349.8	34.1	612	6	CD813712 EN15.020L
16	330.2	32.2	785	8	BZ061279 lki10h12
17	329.2	32.1	721	8	BH969636 odd94e04
18	317.6	30.9	834	7	CO108208 GR_Eb003
19	317.3	30.9	317	1	AV526843 AV526843
20	308.6	30.0	669	1	AV821829 AV821829
21	301.4	29.3	614	1	AV539303 AV539303
22	283.2	27.6	834	7	CO094348 GR_Eal6G
23	273.2	26.6	726	5	EX835627 EX835627
24	267.6	26.1	841	7	CO132469 GR_Eb45F

25	255	24.8	809	7	CO099865
26	255	24.8	872	7	CO128357
27	246.6	24.0	591	6	CA781580 024G12AF
28	236	23.0	453	2	AW705298 ek59c02.v
c 29	234.8	22.9	821	7	CO108207 GR_Eb003
c 30	234.8	22.9	822	7	CO108219
c 31	231.6	22.6	712	7	CO132468 GR_Eb45F
c 32	227.6	22.2	471	1	AI992821 701493711
33	225.4	21.9	870	4	BM360659 GA_Ea003
34	225	21.9	458	7	T21005 3013.Lambda
35	224.8	21.9	748	7	CO084070 GR_Ea49H
36	224.6	21.9	537	4	BM143791 gaJ48910.
c 37	224	21.8	421	1	AV530494 AV530494
38	221.6	21.6	636	6	CD038665
39	221.2	21.5	582	2	AW187216 BNLGH112
40	219.8	21.4	617	4	BM814806 EST592900
41	219.8	21.4	631	1	AJ504308 AJ504308
42	216.6	21.1	624	2	AW54452 N210302e
43	216.2	21.1	668	5	BQ402703 GA_Ed005
44	216.2	21.1	690	4	BG441541 GA_Ea001
c 45	216	21.0	414	1	AV533100 AV533100

ALIGNMENTS

RESULT 1
CNS0A70B
LOCUS
DEFINITION
CNS0A70B
1307 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTFB502E05 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
ACCESSION
BX822585
VERSION
HTC; GSFT_cDNA
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 1307)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
TITLE
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1307)
Genoscope.
JOURNAL
Direct Submission
AUTHORS
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
Location/Qualifiers
1. 1307
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"

FEATURES

source
1. 1307
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"

QY 61 AGGTCCTCCAGTAAGTTTCTGAGTAATTTCAAGGATGAGCTTTCGGTTCCTGCTGACATC 120
Db |||||||
QY 292 AGGTCCTCCAGTAAGTTTCTGAGTAATTTCAAGGATGAGCTTTCGGTTCCTGCTGACATC 351
Db |||||||
QY 121 ACAGTACCTTGATGGTCAATGTTTGGCGGTGAGGACTAAGAAAGCTGACAAACAAATTTGG 180
Db |||||||
QY 352 ACAGTACCTTGATGGTCAATGTTTGGCGGTGAGGACTAAGAAAGCTGACAAACAAATTTGG 411
Db |||||||
QY 181 TTTCAAGATGTTTGGCAAGATTTGTTGACCGTTACTCCTCAATCGCATTTGTTATCTTTTG 240
Db |||||||
QY 412 TTTCAAGATGTTTGGCAAGATTTGTTGACCGTTACTCCTCAATCGCATTTGTTATCTTTTG 471
Db |||||||
QY 241 ATTTTATAGATATCAAGAACTCTGCTTCAGGCTCAATTTCAATTTATCCACTCT 300
Db |||||||
QY 472 ATTTTATAGATATCAAGAACTCTGCTTCAGGCTCAATTTCAATTTATCCACTCT 531
Db |||||||
QY 301 GAGATCAATTTACCATTTCCACCGTCTCATGGATTCGGCTCACACACCTTTCAAGCGGC 360
Db |||||||
QY 532 GAGATCAATTTACCATTTCCACCGTCTCATGGATTCGGCTCACACACCTTTCAAGCGGC 591
Db |||||||
QY 361 CGTTTGTGTTGAAGACCTTGAAGATGAAGATGCGGAGTCACTCTTCTTCTTCTGTGATC 420
Db |||||||
QY 592 CGTTTGTGTTGAAGACCTTGAAGATGAAGATGCGGAGTCACTCTTCTTCTTCTGTGATC 651
Db |||||||
QY 421 CCATCACCACTTCTGAGTCTACAGTACAGCCCAACAAAGGATGCTAGTTTCAAGCATC 480
Db |||||||
QY 652 CCATCACCACTTCTGAGTCTACAGTACAGCCCAACAAAGGATGCTAGTTTCAAGCATC 711
Db |||||||
QY 481 CAACTCTGTTCACTGGACCTTAAAGCTGAAGGAGCCCAACCCCAACCAAAATACCT 540
Db |||||||
QY 712 CAACTCTGTTCACTGGACCTTAAAGCTGAAGGAGCCCAACCCCAACCAAAATACCT 771
Db |||||||
QY 541 AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db |||||||
QY 772 AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 830
Db |||||||
QY 601 CGAGATGATGATCCAGAGAACCGTTCAAAAGTTCTACGAGAGTCTTCTCGAGAAAGAGA 660
Db |||||||
QY 831 CGAGATGATGATCCAGAGAACCGTTCAAAAGTTCTACGAGAGTCTTCTCGAGAAAGAGA 890
Db |||||||
QY 661 ACCGTGACTGCAAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Db |||||||
QY 891 ACCGTGACTGCAAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 950
Db |||||||
QY 721 AACCTCTTCTTCAAGAGTGGTCTGCGACCATCTATCTATACAGAGGTTGCAATCATGAT 780
Db |||||||
QY 951 AACCTCTTCTTCAAGAGTGGTCTGCGACCATCTATCTATACAGAGGTTGCAATCATGAT 1010
Db |||||||
QY 781 CTTCTCTTCTGTTGCTGAGAGTACCTTAAGTGGGATCTCCGGTTCATCAAGTCCAG 840
Db |||||||
QY 1011 CTTCTCTTCTGTTGCTGAGAGTACCTTAAGTGGGATCTCCGGTTCATCAAGTCCAG 1070
Db |||||||
QY 841 CTTGGGAGAGAAATGGCTGTTGATGCTCTACAAAGCCGGGAGGAGGAGGAGGAGGAGGAG 900
Db |||||||
QY 1071 CTTGGGAGAGAAATGGCTGTTGATGCTCTACAAAGCCGGGAGGAGGAGGAGGAGGAGGAG 1130
Db |||||||
QY 901 CAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db |||||||
QY 1131 CAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1190
Db |||||||
QY 961 GAGCTGCTCAGAACACAGAGATTTCTGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020
Db |||||||
QY 1191 GAGCTGCTCAGAACACAGAGATTTCTGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1250
Db |||||||
QY 1021 GTCTGAA 1027
Db |||||||
QY 1251 GTCTGAA 1257
Db |||||||

RESULT 3

CNS0A6S0

LOCUS

DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone

CNS0A6S0

1375 bp

mRNA

linear

HTC 07-FEB-2004

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
Location/Qualifiers
source

gene

ORIGIN

Query Match
Best Local Similarity
Matches 1008; Conservative

97.0%; Score 996.6; DB 3; Length 1375;
98.1%; Pred. No. 1.8e-288;
0; Mismatches 19; Indels 0; Gaps 0;

QY 1 ATGCCACGCCCTTCTCCATAAGTTGATTTTCTCATCCACTATCCAGAAAAAAGCTCTG 60
Db |||||||
QY 225 ATGCCACGCCCTTCTCCATAAGTTGATTTTCTCATCCACTATCCAGAGGAGGAGCTCTG 284
Db |||||||
QY 61 AGGTTCCAGATAGTATTTGTGAGTAAATTCAGAGATGAGCTTTCGGTGTCTGTGCACTC 120
Db |||||||
QY 285 ATGTTCCAGATAGTATTTGTGAGTAAATTCAGAGATGAGCTTTCGGTGTCTGTGCACTC 344
Db |||||||
QY 121 ACAGTACCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 180
Db |||||||
QY 345 ACAGTACCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 404
Db |||||||
QY 181 TTTCAAGATGTTGGCAAGATTTGTTGACCGTTACTCCTCAATCGCATTTGTTATCTTTTG 240
Db |||||||
QY 405 TTTCAAGATGTTGGCAAGATTTGTTGACCGTTACTCCTCAATCGCATTTGTTATCTTTTG 464
Db |||||||
QY 241 ATTTTATAGATATGAAGAACTCTGCCCTTCAGCGCTACATTTTCAATTTATCCACTCT 300
Db |||||||
QY 465 ATTTTATAGATATGAAGAAATTTCTGCCCTTCAGCGCTACATTTTCAATTTATCCACTCT 524
Db |||||||

GLSTPGH542B04 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
BX824523
GI:42465662
HTC; GLST cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1375)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1375)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?ggb?source=Arabidopsis.

Location/Qualifiers
1..1375
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GLSTPGH542B04"
/tissue type="Hormone Treated Callus"
/plasmid="pCMVSPORT 6"
complement(1..1375)
/gene="At3g18990"


```
QY 301 GAGATCAATTACCAATTCACCGGCTCATGATTCGGCTCAACACCACTTCACAAACGCGCC 360
DB 525 GAGATCAATTACCAATTCACCGGCTCATGATTCGGCTCAACACCACTTCACAAACGCGCC 584
QY 361 CGTTGTTTGAAGACCTTGAAGATGAAGATGCGAGGTCATCTTCCCTTCTCTGTGTAC 420
DB 585 CGTTGTTTGAAGACCTTGAAGATGAAGATGCGAGGTCATCTTCCCTTCTCTGTGTAC 644
QY 421 CCATCACCACCTTCCTGAGTCTACACTACAGCCCAACAAAGGTATGCTAGTTCAGCCATC 480
DB 645 CCATCACCACCTTCCTGAGTCTACACTACAGCCCAACAAAGGTATGCTAGTTCAGCCATC 704
QY 481 CAAACCTTGTTCACCTGACCACTTAAAGCTGAAGAGCCAAACCCCAACCCCAAAATACCT 540
DB 705 CAAACCTTGTTCACCTGACCACTTAAAGCTGAAGAGCCAAACCCCAACCCCAAAATACCT 764
QY 541 AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 765 AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 824
QY 601 CGAGATGATGATCCAGAGAACCGTTCAAAGTCTACGAGAGTCTTCTGCGAGAAAGAGA 660
DB 825 CGAGATGATGATCCAGAGAACCGTTCAAAGTCTACGAGAGTCTTCTGCGAGAAAGAGA 884
QY 661 ACCGTGACCTGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
DB 885 ACCGTGACCTGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 944
QY 721 AACCTTCTTCAGAGTGTCTGCGACATCTATCTATACAGAGGTTGATCATGTAT 780
DB 945 AACCTTCTTCAGAGTGTCTGCGACATCTATCTATACAGAGGTTGATCATGTAT 1004
QY 781 CTTCTCTCTGGGTTGCTGAGAGTACCTAAGTGGGATCTCGGCTTCATCAAGTCCAG 840
DB 1005 CTTCTCTCTGGGTTGCTGAGAGTACCTAAGTGGGATCTCGGCTTCATCAAGTCCAG 1064
QY 841 CTTGCGGAGAAACAAATGGCTCTTCGATGCTCTACAAAGCGGAGAGGAGGAGGAGGAGGAG 900
DB 1065 CTTGCGGAGAAACAAATGGCTCTTCGATGCTCTACAAAGCGGAGAGGAGGAGGAGGAGGAG 1124
QY 901 CAAGATGTTAGCAATTCATCTAGAGAACAACTTAGAGAGAGGAGAGGAGGAGGAGGAGGAG 960
DB 1125 CAAGATGTTAGCAATTCATCTAGAGAACAACTTAGAGAGAGGAGAGGAGGAGGAGGAGGAG 1184
QY 961 GAGCTGCTCAGAACAGAGATTTGCTTTTGAAGTGACAGGCTTTCGAGTCAACAGGTAC 1020
DB 1185 GAGCTGCTCAGAACAGAGATTTGCTTTTGAAGTGACAGGCTTTCGAGTCAACAGGTAC 1244
QY 1021 GTCTGAA 1027
DB 1245 GTCTGAA 1251
```

```
RESULT 4
LOCUS CD822418
DEFINITION BN25.045C06F020108 BN25 Brassica napus cDNA clone BN25045C06, mRNA
sequence.
ACCESSION CD822418
VERSION CD822418.1 GI:32504358
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
REFERENCE 1 (bases 1 to 701)
AUTHORS Genoplante.
TITLE Arabidopsis thaliana Full-length cDNA Complete sequence from clone
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
```

93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>)
and <http://genoplante-info.infobiogen.fr>.

FEATURES

source
1..701
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet Neuf"
/db_xref="taxon:3708"
/clone="BN25045C06"
/tissue_type="seed"
/clone_id="BN25"

ORIGIN

Query Match 45.8%; Score 470.8; DB 6; Length 701;
Best Local Similarity 88.0%; Pred. No. 4.3e-130;
Matches 537; Conservative 0; Mismatches 67; Indels 6; Gaps 2;
QY 1 ATGCCAGCCCTTCTTCCATAAGTTGATTTTCTCATCCACTATCCAAAGAAAAGCTGTG 60
DB 92 ATGCCAGCCCTTCTTCCACAGCTGATTTTCTCATCCACTATCCAAAGAAAAGCTGTG 151
QY 61 AGGGTCCAGATAAGTTTGTGAGTAAATTCAGGATGAGCTTTCGGTTGCTGTTCACCTC 120
DB 152 AGAGTTCCTGATAGTTTGTGAGTAAATTCAGGACGAGCTATCGTTGCGGTTGCTCTC 211
QY 121 ACAGTACCTGATGTCATGTTTGGCGTGTAGACTAGGAAAAGCTG---ACAACAATAAT 177
DB 212 ACAGTACCTGATGTCATGTTTGGCGTGTAGACTAAGGAAAAGCTGACACACACAAAATC 271
QY 178 TGGTTTCAAGATGTTTGGCAAGAGTTTGTTCACCGTTACTCCATTTCGATTGGTTATCTT 237
DB 272 TGGTTTCAAGATGTTTGGCAAGAGTTTGTTCACCGTTACTCCATTTCGATTGGTTATCTT 331
QY 238 TTGATTTTATGATGAGGAAACCTCTGCTTCAGCGCTACATTTTCAATTTATTCACAC 297
DB 332 TTGATTTTATGATGAGGAAACCTCTGCTTCAGCGCTACATTTTCAATTTATTCACAC 391
QY 298 TCTGATGATCAATTAACCATTCACCGGCTCATGATTCGCTCAACACCTTCACAAAGC 357
DB 392 TCCGCGATCAACTACCACTCCACCGGCTCATGATTCGACACCAACCACTTCAACAGT 451
QY 358 GCCGTTTGTGTTGAAGACCTTGAAGATGAAGATGCGAGGTCATCTTTCCTTCTCTGTG 417
DB 452 GCGGTTTGTGTTGAAGACCTTGAAGATGAAGATGTTGAGGTTGTTCAACCTTCTCTTGG 511
QY 418 TACCCATCACCACCTTCTGAGTCTACAGTACAGCCAAACAAAGGATGCTAGTTTCAGCC 477
DB 512 TACCCATCAGAGCATCTCTGAGACTACTGCGCACGCTAATAAAGGCGATAGTTTCAGT 571
QY 478 ATCCAAACCTTGTTCACCTGGACCA---GTTAAAGCTGAAGCCACCCACCCCAAAA 534
DB 572 ATCCAGAGCTTCTTCCGCTGAACCCACCTGTTAAAGCTGAAGAGACACACCCCAAAA 631
QY 535 ATACTAAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 594
DB 632 GTTCTTAAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 691
QY 595 GCTCCGCGAG 604
DB 692 GCTCCAGAG 701

RESULT 5

CNSOAD3Z

LOCUS

DEFINITION

ACCESSION

CNSOAD3Z
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTPGH45ZC05 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (Chale cress).
BX816391

VERSION BX816391.1 GI:42473554
 KEYWORDS HTC; GSLT cDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 1420)
 AUTHORS Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
 TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1420)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
 COMMENT The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
 URV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.
 FEATURES
 source
 1..1420
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLTPGH452C05"
 /tissue_type="Hormone Treated Callus"
 /plasmid="pCMVSPORT 6"
 complement(1..1420)
 /gene="At1g49480"
 gene
 ORIGIN
 Query Match 42.9%; Score 441; DB 3; Length 1420;
 Best Local Similarity 73.5%; Pred. No. 5.4e-121;
 Matches 646; Conservative 0; Mismatches 195; Indels 38; Gaps 5;
 QY 151 GGACTAAGGAAAGCTGACAAACAAATTTGGTTTCAAGATGGTTGGCAAGAGTTCTTCTGAC 210
 DB 336 GGCATGAAGAAAGCCAAACAAACAAATCTGTTTCAAGACGGTTGGCAGAGTTTGTCAAC 395
 QY 211 CGTTACTCATTCGCGATTGGTTATCTTTGATTTTATAGATGAAGGAACTCTCCCTTC 270
 DB 396 CGTTTCTCATTCGCGATTGGT-----TTGATGATCAAGATTACAGTCTACATTTCAATTT 451
 QY 271 AGCGTCTCAATTTCAATTTATTCACACTCTGAGTCAATATACCATTCACCGGTTCTCATG 330
 DB 452 ATCTCTCCACACTCTGAGATCAACACC-----ATTCTAGTAGTGAAGCTCTTATGCAATG 507
 QY 331 GATTCGGCTCAACACACTT-----CAACGCGCGCTTTGTTTGAAGACCTTGAAGATGAA 387
 DB 508 GATTCGGCAGCAATCAGTTCAACAAACGTCGCTCGATTGTTTGAAGATCTCTGAATCAAA 567
 QY 388 GATGCGGAGGTGATCTTTCTCTTCTCTGTTGATCCCATCACCATTCTCTGAGTCTACGTA 447
 DB 568 GATGCTAAGTCAATTTATC-----CATCGAACCCCTGAATCTACTGAA 609
 QY 448 CCAGCCAAACAAGGGTATGCTAGTTGAGCCATCCAAACCTTGTTCATCGGACAGTTAAA 507

DB 610 CCAAGTGAATAAAGGTTATGGCGGTTCTATAGCCATCCAAAGCTTTTTCAAAGAATCTTAAA 669
 QY 508 GCTGAAGAGCAACCGCCCAACCCCAAAAAATACCTAAAGAGAGAGGAGGAGGAAGAAAAT 567
 DB 670 GCTGAAGA-----AACTCCCAAGTACTTAAAGAGAGAGGAGGAGGAAGAAAGAAAT 720
 QY 568 GCTGATCTCTAGGAAATAAATCAATCATAGCTCCGCGAGATGATATCCAGAGAACCGTTCA 627
 DB 721 CTTAATCCGAGGAAGTAAATCTTTCAACTCCCGTGGAGATGACTCAGAGAACCGCTCA 780
 QY 628 AAGTTCTACGAGAGTCTTCTCGGAGGAAGAGAACCGTACTGCGAGAGAAAGAGAGAGA 687
 DB 781 AAGTTCTACGAGAGTCTTCTGCTAGAAAGAGAACTGTAACTGCGAGAGAAAGAGAGAGA 840
 QY 688 GCCATCAATGAGCCCAAAACGTTTCCAAACCAACCCCTTTCTTCAGAGTGGTTCTGCGA 747
 DB 841 GCCGTCATGAGCCCAAAACATTCGAACCAACAACTCTTACTTTAGAGTTGTTCTGCGA 900
 QY 748 CCATCTATCTATACAGAGTTGCAATCATGATCTTCTCTCTGCTGGTTGCTGAGAAATAC 807
 DB 901 CCATCATATCTATACAGAGTTGCAATCATGATCTTCTCTCTGCTGGTTGCTGAGAAATAC 960
 QY 808 CTAGTGGATCTCCGGTTTCATCAAACTCCAGCTTCGGGAGAAACAATGGCCTTTCGA 867
 DB 961 CTAAGTGGATATCTGGTTTCATCAAGCTCCAGCTCGGTGAGAAACAATGGCCAGTGG 1020
 QY 868 TGTCTCTACAAAGCGCGAGAGCCAAATTCAGTCAAGGATGATGATGATTCATCTACATAG 927
 DB 1021 TGCCTCTACAAAGCGAGAGCTAAGTTTAGCCAGGATGGTATGATGATGATGATGATG 1080
 QY 928 AACCACTTAGGAGAGGAGAGCTGCTGCTGTTTGAAGTCTGCTGAGAACGAGAGATTCGTT 987
 DB 1081 AACCAATATAGCGAAGGAGATGATGTGTGTTGAGTACTCAGAACTCGGGATTCGTT 1140
 QY 988 TTGAAGTGCAGAGCTTTCGAGTCAACGAGTACGTTCTGA 1026
 DB 1141 CTCGAAGTCAACCGCTTTTCGTTCAATGATGATGTTGA 1179
 RESULT 6
 CDSAD86 1490 bp mRNA linear HTC 06-FEB-2004
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 DEFINITION GSUTPGH132B06 of Hormone Treated Callus of strain col-0 of
 Arabidopsis thaliana (thale cress).
 ACCESSION BX815859
 VERSION BX815859.1 GI:42473387
 KEYWORDS HTC; GSLT cDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 1490)
 AUTHORS Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
 TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1490)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
 COMMENT The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
<http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis>.

FEATURES

Location/Qualifiers
1..1490
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTPGH132B06"
/tissue_type="Hormone Treated Callus"
/plasmid="pCMVSPORT 6"
complement(1..1490)
/gene="At1g49480"

ORIGIN

Query Match 42.9%; Score 441; DB 3; Length 1490;
Best Local Similarity 73.5%; Pred. No. 5.5e-121;
Matches 646; Conservative 0; Mismatches 195; Indels 38; Gaps 5;
QY 151 GCACTAAGGAAGCTGACACAAATTTGGTTTCAAGATGGTTGCAAGAGTTTCTGTGAC 210
DB 225 GACTAAGGAAGCCACACAAATCTGGTTTCAAGACGGTTGCGAGAGTTTGTCAAC 284
QY 211 CGTTACTCCATTCGATGGTTTATCTTTGATTTTATGATATGAAGGAACTTCGCTTC 270
DB 285 CGTTTCTCCATTCGGATGGT---TTTCAGATACAAAGTACAGTCTACATTTCAAATTT 340
QY 271 AGCGTCTACATTTCAATTTATCCACTCTGAGATCAATTACCATTCCACCGTCTCATG 330
DB 341 ATCTCCACACTCTGAGATCAACACC---ATTCTAGTAGGAAGCTCTTATGCAATG 396
QY 331 GATTCGGCTCACACACACTT---CAACGCGCGGTTTGTGGAAGACCTTGAAGATGAA 387
DB 397 GATTCGCGACAGATCAGTTCAACAAAGTGTCTGATTTGATGATCTTGAATCTGAACTCAA 456
QY 388 GATCGCGAGGTCAATCTTCTCTTCTGTGTATCCCATCACTTCTCTGAGTCTACATGA 447
DB 457 GATCGGAAGGTCAATTC---CATCGAACCTTGAATCTACTGAA 498
QY 448 CCAGCCACAAAGGTATCTATGTTTCAGCCATCCAAACCTTGTTCCTGACCACTTAA 507
DB 499 CCAGTGAATAAAGGTTATGGCGGTTCTACAGCCATCCAAAGCTTTTCAAAGAATCTAAA 558
QY 508 GCTGAAGAGCCAAAGCCAAACCCAAATAATCTTAAAGAGAGGAGGAGGAAGAAAT 567
DB 559 GCTGAAGA-----AAGCGCCAGGTACTTTAAGAGAGAGGAGGAGGAGGAGGAGAT 609
QY 568 GCTGATCTCTAGGAATAAATCATCATCGCTCCGCGAGATGATGATCCAGAGAACCGTTCA 627
DB 610 CCTAATCCCGAGGAAGTAATCTTCAACTCCCGGTGGAGATGATCTCAGAGAACCGCTCA 669
QY 628 AGTTCTACGAGAGTCTTCTCGGAGAGAGAGAACCGTGACTGCGAGAGAGAGAGAGA 687
DB 670 AAGTCTACGAGAGTCTTCTGCTGAGAAAGAGAACTGTAATCTCAGAGAGGAAAGAGAGA 729
QY 688 GCATCAATGAGCCAAAGCTTGAACCAACAAACCTTCTTCTCAGAGTGGTTCTGCGA 747
DB 730 GCGTCAATGAGCCAAACATTCGAAACCAATCTTACTTTAGAGTGGTTCTGCGA 789
QY 748 CCATCTATCTATACAGAGGTTGCAATCATGATCTTCTCTTCTGGGTTGCTGAGAGTAC 807
DB 790 CCATCATATCTATACAGAGGTTGCAATCATGATCTTCTGCAATCTGCTGAGAAATAC 849
QY 808 CTAAGTGGATCTCCGGTTCATCAAGTCCAGCTTGGGAGAAACATGGCTGTTCGA 867
DB 850 CTAAGTGGATATCTGGTTTCATCAAGCTCCAGCTCGGTGAGAAACATGGCCAGTGG 909
QY 868 TGTCTCTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGTTGATGATCTACTCTAG 927

DB 910 TGCCTCTACAAAGCAGGAGAGCTAAGTTTAGCCAAGGATGATGAGTTCACACTCGAG 969
QY 928 AACAACTTAGAGAGAGAGAGAGCTGTGTGTGTTTGTAGCTGCTCAGAACAGAGATTTCTTT 987
DB 970 AACAAATAGCGGAGGAGATGATGTGTGTTTGTAGCTACTCAGAACTCGGATTTCTGTT 1029
QY 988 TTGAAGTGACAGCTTTTCGAGTCAACGAGTACGTCTGA 1026
DB 1030 CTCGAAGTCACCGCTTTCGTGTCATGATGTGTGA 1068
CNS0AD6P 1346 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTPGH542B06 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
BX816536
ACCESSION
VERSION BX816536.1 GI:42473594
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1346)
CASTELLI,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1346)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
<http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis>.
Location/Qualifiers
1..1346
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTPGH542B06"
/tissue_type="Hormone Treated Callus"
/plasmid="pCMVSPORT 6"
complement(1..1346)
/gene="At1g49480"

FEATURES

source

gene

ORIGIN

Query Match 42.6%; Score 437.8; DB 3; Length 1346;
Best Local Similarity 73.3%; Pred. No. 4.8e-120;
Matches 644; Conservative 0; Mismatches 197; Indels 38; Gaps 5;
QY 151 GCACTAAGGAAGCTGACACAAATTTGGTTTCAAGATGGTTGCAAGAGTTTCTGTGAC 210
DB 229 GCACTAAGGAAGCCAAACAAATCTGTTTCAAGACGGTTGCGAGAGTTTGTCAAC 288


```
Db 386 CCTAATCCGAGGAAGTAACTTTCAATCCCGTGGAGATGACTCAGAGAACCGCTCA 445
QY 628 AAGTTCTACGAGAGTCTTCTCGGAGAAAGAGAACCGTACTGC-AGAAGAAAGAGAG 686
Db 446 AAGTTCTACGAGAGTCTTCTCGGAGAAAGAGAACCGTACTGC-AGAAGAAAGAGAG 505
QY 687 AGCCATCAATGACGACAAACGTTTCGAAACCAACCAACCTTTCTTTCAGAGTGTCTCGG 746
Db 506 AGCGCTCAATGACG-CAAAACATTGCAACCAACCAATCTTACTTTAGAGTGTCTCGG 564
QY 747 ACATCTCTATCATACAGAGTGTGATCATGATCTTCTTCTCGGTTTTCGAGAGTA 806
Db 565 ACCATCATATCTATACAGAGTGTGATCATGATCTTCTTCTCGGTTTTCGAGAGTA 624
QY 807 CCTAAGTGGGATCTCGGGTTCATCAAGTCCAGTTCGGGAGAAACCAATGSCCTGTCG 866
Db 625 CCTAAGTGGGATCTCGGGTTCATCAAGTCCAGTTCGGGAGAAACCAATGSCCTGTCG 684
QY 867 ATGTCTCTACAAAGCGGGAGAGCAAAATTCAGTCAAGGATGTTACGAAATTCACCTCTAGA 926
Db 685 GTGCTCTACAAAGCGGGAGAGCAAAATTCAGTCAAGGATGTTACGAAATTCACCTCTAGA 744
QY 927 GAACAACTTAGAGAGAGGAGAGTCTGTGTGTTTTCAGTCTGCTCAGAACCAAGATTTGCT 986
Db 745 GAACGATATAGCGAGGAGAGATGATGTGTGTTTTCAGTCTGCTCAGAACCAAGATTTGCT 804
QY 987 TTTGAAAGTGACAGCTTTTCGAGTCAACGAGTACGCTCTGA 1026
Db 805 TCTCGAAGTCCCGCTTTTCGTTCAATGATGATGTGTA 844
```

```
RESULT 9
AV540912/c
LOCUS AV540912 589 bp mRNA linear EST 20-FEB-2004
DEFINITION AV540912 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
CDNA clone RZ157509F 3', mRNA sequence.
ACCESSION AV540912
VERSION AV540912.1 GI:8702670
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 589)
REFERENCE Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
AUTHORS A large scale analysis of cDNA in Arabidopsis thaliana: Generation
TITLE of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..589
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="RZ157509F"
/tissue_type="roots"
/clone_lib="Arabidopsis thaliana roots Columbia"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI"
```

FEATURES

source

Query Match 39.1%; Score 401.8; DB 1; Length 589;
Best Local Similarity 99.5%; Pred. No. 2.5e-109;
Matches 403; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN

Query Match 39.1%; Score 401.8; DB 1; Length 589;
Best Local Similarity 99.5%; Pred. No. 2.5e-109;
Matches 403; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 623 GTTCAAAGTTCTTACGAGAGTCTTCTCGGAGAAAGAGAACCGTACTGCAGAGAAAGAG 682
Db 589 GTTCAAAGTTCTTACGAGAGTCTTCTCGGAGAAAGAGAACCGTACTGCAGAGAAAGAG 530
QY 683 AGAGAGCCATCAATGACGACCAAAACGTTTCGAAACCAACCAACCTTTCTTTCAGAGTGTTC 742
Db 529 AGAGAGCCATCAATGACGACCAAAACGTTTCGAAACCAACCAACCTTTCTTTCAGAGTGTTC 470
QY 743 TCGGACCAATCTATCTATACAGAGTGTGATCATGATCTTCTTCTCGGTTTTCGCTGAGA 802
Db 469 TCGGACCAATCTATCTATACAGAGTGTGATCATGATCTTCTTCTCGGTTTTCGCTGAGA 410
QY 803 AGTACTAAGTGGGATCTCCGGTTCATCAAGTCCAGTTCGGGAGAAACCAATGSCCTG 862
Db 409 AGTACTAAGTGGGATCTCCGGTTCATCAAGTCCAGTTCGGGAGAAACCAATGSCCTG 350
QY 863 TTGATGTTCTTCAAAAGCGGGAGAGCAAAATTCAGTCAAGGATGTTTCAGAAATTCATC 922
Db 349 TTGATGTTCTTCAAAAGCGGGAGAGCAAAATTCAGTCAAGGATGTTTCAGAAATTCATC 290
QY 923 TAGAGAACAACTTAGGAGAGGAGACCTCTGTGTGTTTTCAGTCTCAGAACCAAGATTT 982
Db 289 TAGAGAACAACTTAGGAGAGGAGACCTCTGTGTGTTTTCAGTCTCAGAACCAAGATTT 230
QY 983 TCGTTTTCGAAAGTGACAGCTTTTCGAGTCAACGAGTACGCTCTGAA 1027
Db 229 TCGTTTTCGAAAGTGACAGCTTTTCGAGTCAACGAGTACGCTCTGAA 185
```

RESULT 10

CK120978

LOCUS

CK120978

DEFINITION

CK120978

ACCESSION

CK120978

VERSION

CK120978.1

KEYWORDS

EST.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 780)

REFERENCE Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.

AUTHORS Generation of a cDNA expression library from Arabidopsis

inflorescence meristem

Unpublished (2003)

Contact: Birgit Kersten

Plant Protein Chip Group, Department Leirach

Max-Planck-Institute for Molecular Genetics

Inhestr. 73, D-14195 Berlin, Germany

Tel: +49(0)30/84131648

Fax: +49(0)30/84131128

Email: Kersten@molgen.mpg.de

Insert Length: 780 Std Error: 0.00

Plate: 204 row: 0 column: 14

Seq primer: PQ85

Location/Qualifiers

1..780

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/ecotype="Columbia"

/db_xref="taxon:3702"

/db_xref="GABI:952488"

/db_xref="taxon:3702"

/clone="MPMGp2011014204"

/tissue_type="inflorescence meristem"

/dev stages="about one week after bolting"

/lab_host="E. coli SCS-1/pSE111"

/clone_lib="AtM1"

/note="Vector: pQE-30NAST-attB (AY386205); Site 1: Salt;

Site 2: NotI; About 1 week after bolting, cDNA synthesis

using SuperscriptTM-system (Invitrogen) with an

oligo(dT)-primer containing NotI restriction site and a
Sali adapter. The main library (plate numbers begin with
1) of 38,000 clones was rearranged into the sublibrary
(plate numbers begin with 201) containing 5,000 putative
expression clones. Average insert size is 1 kb. Note: The
rearranged sublibrary (plate numbers begin with 201) was
sequenced. Library generation and sequencing was granted
in context of GABI-LAPP; data are also accessible at
<https://gabi.rzd.de>

ORIGIN

Query Match 38.6%; Score 396.2; DB 7; Length 780;
Best Local Similarity 72.9%; Pred. No. 1.4e-107;
Matches 594; Conservative 0; Mismatches 183; Indels 38; Gaps 5;

```
QY 159 GAAAGCTGACACAAATTTGGTTTCAAGATGGTGGCAAGAGTTTGTGACCGTTACTC 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GAAAGCCACACAAATTTGGTTTCAAGATGGTGGCAAGAGTTTGTGACCGTTACTC 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 219 CATTCGCATGTGTTATCTTTTGAATTTTATAGATATGAAGAAACTCTGCCCTTCAGCGTCTA 278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 CATTCGCATGTGTTATCTTTTGAATTTTATAGATATGAAGAAACTCTGCCCTTCAGCGTCTA 116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 279 CATTTTCATTTATCCACTCTGAGATCAATTACCAATTCACCGGTCTCATGATTCGCC 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 117 CACTCTGAGATCAACACC-----ATTCTAGTAGTGAAGCTCTTATGCAAAATGGATTCGCC 172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 339 TCACAAACCACTT---CAAAGCGCGCGTTTGTGTTGAAGACCTTGAAGATGAAGATGCCGA 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 ACAGATCATGTTCAACAAAGCTGCTCATGTTTGAAGATCTGAACTCAAGATGCTAA 232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 396 GGTCACTTTTCCCTTCTTCTGTGTACCATCACCACCTTCTGAGTCTACAGTACCAGCCAA 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 GGTCACTTTATC-----CATCGAACCTGTAATCTACTGTAACCACTGAA 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 456 CAAAGGATGTAGTTTACGCCATCCAAACCTTTTCACTGGACGCTTAAAGCTGGA 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 TAAAGGTTATGGCGTTCTACAGCCATCCAAAGCTTTTCAAGATCTTAAAGCTGAAGA 334
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 516 GCCAACGCCAACCCCAAAATACCTAAAGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 -----AACGCCCAAGTACTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 576 TGAGGAAATAAACTCATAGCTCCGCGAGATGATGATCAGAGAACCGTTCAAAGTTCTA 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 CGAGGAAGTAAACTCTTCAACTCCCGGTGAGATGACTCAGAGAACCGTCAAAGTTCTA 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 636 CGAGAGTGTCTTCGAGAGAAAGAAACCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 446 CGAGAGTGTCTTCTGTAGAAAGAGAACTGTAACTGCAGAGAGAGAGAGAGAGAGAGAGAG 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 696 TGCAGCCAAACGTTCCGAACCAACCAACCTTTCTTCAGAGTGGTTCTCGGACCACTCTA 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 TGAGCCAAACCAATTCGAACCAACCAACCTTTCTTCAGAGTGGTTCTCGGACCACTCTA 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 756 TCTATACAGAGGTTGCATCATGTATCTTCTCTTCTGGGTTTGTCTGAGAAGTACTAAGTGG 815
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 TCTATACAGAGGTTGCATCATGTATCTTCTCTTCTGGGTTTGTCTGAGAAGTACTAAGTGG 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 816 GATCTCCGGTTTCATCAAGTCCAGTTCGGAGAGAAACAAATGGCTGTTCGATGTCTCTA 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 626 GATATCTGGTTTCATCAAGCTCCAGTTCGGTGAAGAAACAAATGGCCAGTGGGTGCTCTA 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 876 CAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 CAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 936 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 746 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 11

A1996485/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .591

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/ecotype="Col-0"

/db_xref="taxon:3702"

/clone="701666887"

/tissue_type="root"

/dev_stages="4 - 7 weeks"

/clone_lib="A. thaliana, Columbia Col-0, root-1"

/note="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA library was derived from untreated root tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunted, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

the pSPORT vector."

Query Match 37.7%; Score 387; DB 1; Length 591;

Best Local Similarity 99.5%; Pred. No. 7.4e-105;

Matches 409; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 618 GAACCGTTCAAAGTTCTACGAGAGTCTTCGCGAGAAAGAACCG-TGACTGCAGAG 676

Db 591 GAACCGTTCAAAGTTCTACGAGAGTCTTCGCGAGAAAGAACCGTTGATGCAGAG 533

QY 677 AAAG 736

Db 532 AAAG 473

QY 737 TGGTCTCGGACCATCTATCTATACAGAGGTTGCATCATGTCCTCTCTCTCGGTTTG 796

Db 472 TGGTCTCGGACCATCTATCTATACAGAGGTTGCATCATGTCCTCTCTCTCGGTTTG 413

QY 797 CTGAGAAGTACTAAGTGGATCTCCGGTTCATCAAAAGTCCAGCTTTCGGAGAAACAAT 856

Db 412 CTGAGAAGTACTAAGTGGATCTCCGGTTCATCAAAAGTCCAGCTTTCGGAGAAACAAT 353

QY 857 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 916

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db 916 GGCCTGTTTGCATGTCCTTACAAAGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

Db	352	GGCGTGTTCGATGCTCTTCAAAAGCCGGGAGAGCCAAATTCAGTCAAGATGGTACGAAT	293
Qy	917	TCACCTCTAGAGAAACAACCTTAGGAGAGGAGACGCTCTGTGTTTGGAGTGTCTCAGAAACCA	976
Db	292	TCACCTCTTAGAGAACAACTTAGGAGAGGAGACGCTCTGTGTTTGGAGTGTCTCAGAAACCA	233
Qy	977	GAGATTTTCGTTTTTGAAGTGACAGCCTTTCGAGTCAACGAGTACGTCTGAA	1027
Db	232	GAGATTTTCGTTTTTGAAGTGACAGCCTTTCGAGTCAACGAGTACGTCTGAA	182
RESULT 12	EX835682/c		
LOCUS		1041 bp mRNA linear EST 11-FEB-2004	
DEFINITION		EX835682 Arabidopsis thaliana Adult vegetative tissue Col-0	
		Arabidopsis thaliana cDNA clone GSLTULS482H11 3PRIM, mRNA sequence.	
ACCESSION		EX835682	
VERSION		EX835682.1 GI:42529765	
KEYWORDS		EST.	
SOURCE		Arabidopsis thaliana (thale cress)	
ORGANISM		Arabidopsis thaliana	
REFERENCE		Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.	
AUTHORS		Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Coordinated Approach to Evaluate and Improve Arabidopsis Genome	
TITLE			

RESULT 13	
CD823066	
LOCUS	645 bp mRNA linear EST 10-JUL-2003
DEFINITION	BN25.047120F020109 BN25 Brassica napus cDNA clone BN25047120, mRNA sequence.
ACCESSION	CD823066
VERSION	CD823066.1
KEYWORDS	GI:32505006
SOURCE	EST.
ORGANISM	Brassica napus (rape)
REFERENCE	Brassica napus
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
TITLE	1..(bases 1 to 645)
JOURNAL	Genoplatane.
COMMENT	Genoplatane, a major partnership french program in plant genomics Unpublished (2003) Contact: Genoplatane Genoplatane 93, rue Henri Rochefort 91025 EVRY CEDEX France Tel: 33 1 69 47 54 00 Fax: 33 1 69 47 54 10 This sequence has been generated in the framework of the french plant genomics programme 'Genoplatane' (http://www.genoplatane.com and http://genoplatane-info.infobioingen.fr).

Qy	Db	Query Match	Best Local Similarity	Mismatches	Indels	Gaps
434	CTGAGTCTACAGTACCGCCACCAAGAGGTGCTAGTTTCAGCCATCCAAACCTTGTTC	37.3%;	Score 382.6;	DB 5;	Length 1041;	
876	CTGAACACTACTGACCAAGTGAATTAAGGTTATGTGGCGTTCTACGCCATCCAAAGCTTT	79.3%;	Pred. No. 1.9e-103;			
494	CTGGACCACTTAAGCTGGAAGCCACCGCCACCCCAAAAATACCTTAATAAGAGAGGGA	0;	Mismatches 114;			
816	TCAAGAATCTTAAGCTGAAGA-----AACGCCCAAGGTACTTAAGAAGAGAGGAA					
554	GGAGAGAGAAATGCTGATCTTCAGGAATAAATCAATCAGCTCCGCCAGATGATC					

Db 1 GGAGGAGAGAGAGAGAGCGTGTCTGAGAGAGTAACTCATCTGCTCCGAGGGATGATG 60
QY 611 ATCCAGAGAACCGTCTCAAGTCTTACGAGAGTCTCTCGAGAGAGAGAACCGTGACTG 670
Db 61 ACCCGAGAGCGCTTCAAGTCTTACGAGAGTCTCTCGAGAGAGAGAACCGTGTACTG 120
QY 671 CAGAAGAGAGAGAGAGAGCGCATCAATGCGAGCCAAAACGTTGAGAACCAACAAACCCCTTCT 730
Db 121 CAGAGAGAGAGAGAGAGCGCCATTAATGCGAGCCAAAACGTTGAGAACCAACAAACCCCTTCT 180
QY 731 TCAGAGTGTCTCGACACCATCTATCTATACAGAGGTTGCGATCATCTCTCTCTG 790
Db 181 TCAGAGTGTCTCGACACCATCTATCTATACAGAGGTTGCGATCATCTCTCTCTG 240
QY 791 GGTTCCTGAGAGTACCTAAGTGGGATCTCCGGGTTCAATCAAGTCCAGCTTGGCGAGA 850
Db 241 GTTTCCTGAGAGTACTTAAGTGGGATCTCGGGATTCATCAAGTCCAGCTCGGGAGA 300
QY 851 AACAAATGGCTGTTCGATGTCTTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGT 910
Db 301 AACAGTGGCGGTGAGATGCTTTTACAAAGCGGAGAGCCAAATTCAGTCAAGGATGT 360
QY 911 ACGAATTCCTCTAGAGAACCTTAGAGAGAGAGAGAGTCTGTGTGTTGAGTCTCTCA 970
Db 361 ACAGATTCACCTCGAGAGAACCTTAGAGAGAGAGTCTGTGTGTTGAGTCTCTCA 420
QY 971 GAACAGAGATTCGTTTGAAGTGCAGCGCTTTCGAGTCAACG 1015
Db 421 GAACAGAGATTCGTTTGAAGTGCAGCGCTTTCGAGTCAACG 465

RESULT 14

COL08220 GR_Eb0039N17.r GR_Eb Gossypium raimondii cDNA clone GR_Eb0039N17
3', mRNA sequence.

ACCESSION
COL08220
VERSION
COL08220.1 GI:48806906
KEYWORDS
EST.
SOURCE
Gossypium raimondii
ORGANISM
Gossypium raimondii

REFERENCE
AUTHORS
Kim.H., Yu.Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
TITLE
Global assembly of Cotton ESTs
JOURNAL
Unpublished (2004)
COMMENT
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 0039 row: N column: 17.
Location/Qualifiers
1. 906
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb0039N17"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"

FEATURES

source

/note=Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 35.3%; Score 362.6; DB 7; Length 906;
Best Local Similarity 64.8%; Pred. No. 1.9e-97;
Matches 587; Conservative 0; Mismatches 310; Indels 9; Gaps 3;
QY 24 GTTGATTTTCTCATCTATCCAGAAAAAGCTCTGAGGGTCCCAAGATAAGTTTGTGAG 83
Db 1 GCTTATTTCTCTACCACTCCCAAGAAAAAGAACTGAGGATCCCTGATAAATTCGTTAA 60
QY 84 TAAATTTCAAGGATGAGCTTTCGGTGTCTGCTTCCACTCAGTACCTGATGGTTCATCTTG 143
Db 61 GAAATTCAGGATGAGCTTTCGGTGTCTGCTTCCCTCTCACTGTCTCCGCGTCACTGTTG 120
QY 144 GCGTGTAGACTAAGAAAAAGCTGCAACAAAAATTTGGTTTCAAGATGGTTGGCAAGATT 203
Db 121 GGGTGTAGGAATAAAGAAAGTTGACAAAGGTTTGGTTTCAGGAAGGTTGGCAGAGTT 180
QY 204 TGTGACCGTTTACTCCATTCGCAATGGTTATCTTTTGTATTTTGTATGATGAAAGAAATTC 263
Db 181 TCTAGAGCGTTACTATATTCGTGTGGCTACGTATTTGGTTTGTATGATGAAAGAAATTC 240
QY 264 TGCCTTCAGCGTCTACATTTTCAATTTATCCCACTCTGAGATCAATTACCATTCACCGG 323
Db 241 TGCCTTCAGTGTAGTATATTTTGTACAACTCGGAATTAATACTATCAGACTAATGC 300
QY 324 TCTCATGGAATTCGGCTCACAAACACTTCAAAACGCGCCGTTTGTGTAAGAGACCTTCAAGA 383
Db 301 CTTCTGGTGTACTCAATCAATCAGGAACAATATCCATTTGAACAATTCGAGATGA 360
QY 384 TGAAGATGCCAGGTCATCTTTCTTCTGTGTACCATCACCACTTCTCTGAGTCTAC 443
Db 361 TGAATGCATCTCTCCAGCACTTCAGAAATTTGTTTGGTGGTCTAAACTTAAACAATGCAT 420
QY 444 AGTACAGCAACAAAGGATGCT---AGTTGAGCATCCAAACCTTCTCAGTGGACC 500
Db 421 AAACCTGGGGTGTGATGTCATCTTCAAACTTTCAAAAGGTTGTAATAAACAACCTATTTCG 480
QY 501 AGTTAAAGCTGAAGAGCCAAACCCCAAAATATACCTAA---AAAGAGAGGGAGGAA 557
Db 481 AGTCAAAATGTCATCTTCAAGTTTCAAGTGCATGAACCCAGAACCAAAAAGGTTGGAG 540
QY 558 GAAGAAAAATGCTGATCTCTGAGGAATAAATCATCATCAGCTCCGCGAGATGATGATCCAGA 617
Db 541 AAGCGGAGATTTGATCTCTAACTGTCAGGATTCATCTGTCGACGTTGAAGATGATGTTGA 600
QY 618 GAACCGTTCAAAGTTCTACGAGAGTCTTCTCGGAGAAAGAGAACCGTGAATGAGAGAA 677
Db 601 TATGCGCTTTAGATGTTATGAAGTGTCTTCAAGCCCGGAAAGAGAACCGTGAAGTGAAGA 660
QY 678 AAGAGAGAGAGCCATCAATGAGCCAAACCGTTGCAACCAACCAACCCCTTCTTCAGAGT 737
Db 661 AAGAGAGAGAGCAATCAATGAGCCAAACCGTTGAGCCAACTAACCCTTCTTCGAGGGT 720
QY 738 GGTCTTGGGACCATCTATCTATACAGAGTGTGATCATGTATCTTCTTCTGGGTTTGC 797
Db 721 GGTCTTGGGACCATCTATCTGTAAGGGATGATATGATGATCTTCTTCTGCTGCTTTCG 780
QY 798 TGAGAAAGTACCTAAGTGGGATCTCGGGTTCATCAAGTCCAGTTCGCGA---GAAACA 854
Db 781 TGAGAAAGCATCTAAGTGGGTTTCTGGATCCATTAATTAATTTCACTTCTGATGGGAGACA 840
QY 855 ATGGGCTGTTGATGTCCTCAAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGTTACGA 914
Db 841 GTGGTCTGTAAGATGCTTTTATAAGGAGGCAAGGCTAATTTCACTGAGGATGTTATGA 900
QY 915 ATTCAC 920
Db 901 ATTTAC 906

RESULT 15
CD813712

LOCUS CD813712 612 bp mRNA linear EST 10-JUL-2003
DEFINITION BN15.020L13F020211 BN15 Brassica napus cDNA clone BN15020L13, mRNA
sequence..
ACCESSION CD813712
VERSION CD813712.1 GI:32495652
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 612)

AUTHORS Genoplante.

TITLE Genoplante, a major partnership french program in plant genomics

JOURNAL Unpublished (2003)

COMMENT Contact: Genoplante

Genoplante

93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french

plant genomics programme 'Genoplante' (<http://www.genoplante.com>

and <http://genoplante-info.infobiogen.fr>).

FEATURES

Location/Qualifiers

1..612

/organism="Brassica napus"

/mol_type="mRNA"

/cultivar="Jet Neuf"

/db_xref="taxon:3708"

/clone="BN15020L13"

/tissue_type="seed"

/clone_lib="BN15"

ORIGIN

Query Match 34.1%; Score 349.8; DB 6; Length 612;
Best Local Similarity 90.6%; Pred. No. 1.2e-93;
Matches 385; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

QY 1 ATGCCACGCCCTTTCTCCATAAGTTGATTTCTCATCCACTATCCAGAAAAAGCTCTG 60
Db |||||
QY 188 ATGCCACGCCCTTTCTCCACAGCTGATTTCTCATCCACTATCCAGAAAAAGCTCTG 247
Db |||||
QY 61 AGGGTCCCAGATAAGTTGTGAGTAAATTCAGAGATGAGCTTTCGGTTGCTGTGCACCTC 120
Db |||||
QY 248 AGAGTTCCTGATAAGTTGTGAGTAAATTCAGAGGAGAGCTATCGTTGGCGTTGCCCTC 307
Db |||||
QY 121 ACAGTACCTGATGGTCATGTTGGCGGTGAGGACTAAGGAAAGCTG---ACAAACAAAT 177
Db |||||
QY 308 ACAGTACCTGATGGTCATGTTGGCGGTGAGGACTAAGGAAAGCTGACAAACAAATC 367
Db |||||
QY 178 TGGTTTCAAGATGGTTGGCAAGATTTGTTGACCGTTACTCCATTCCGATTGGTTATCTT 237
Db |||||
QY 369 TGGTTTCAAGATGGTTGGCAAGATTTGTTGACCGTTACTCAATCCGATTGGTTACCTT 427
Db |||||
QY 238 TTGATTTTATGATATGAGGAAACTCTGCCTTCAGCGTCTACATTTTCAATTTATCCAC 297
Db |||||
QY 428 TTGATATTATGATACGAGGCACTCTGCTTCAGCGTCTACATTTACAACTTATCACAC 487
Db |||||
QY 298 TCTGAGATCAATATACATTTCCACCGGTCTCATGGAATTCGGCTCAAAACCACTTCAAGCG 357
Db |||||
QY 488 TCCGAGATCAACTACCACTCCACCGGTCTCATGGAATTCGACACAAACCACTTCAAGCGT 547
Db |||||
QY 358 GCCCGTTTGTGAGACCTTGAAGATGAAGATGCGGAGGTCTATCTTTCCTTCTCTGTG 417
Db |||||
QY 548 GCGCGTTTGTGAGACCTCGAAGATGAAGATGTTGAGGTGTTCAACCCCTTCTCCTTTG 607
Db |||||
QY 418 TACCC 422
Db |||||
QY 608 TACCC 612

Search completed: December 30, 2004, 14:32:03

Job time : 3907 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 13:02:12 ; Search time 98 Seconds
(without alignments)
2002.069 Million cell updates/sec

Title: US-10-088-187A-11
Perfect score: 1796
Sequence: 1 MPRPFHKLIFSSTIQEKLRL.....LLRTDRFVLKVTAFRVNEYV 341

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02: *
1: uniprot_spport: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1796	100.0	341	2 Q813W1	Q813W1 arabidopsis
2	1513	84.2	329	2 Q6R2U8	Q6R2U8 brassica ca
3	1513	84.2	329	2 AAR92199	AAR92199 brassica
4	1509	84.0	329	2 Q6V0J6	Q6V0J6 brassica ca
5	1509	84.0	329	2 AAQ55453	AAQ55453 brassica
6	1197	66.6	230	2 Q9LJ67	Q9LJ67 arabidopsis
7	946	52.7	226	2 Q9XIB5	Q9XIB5 arabidopsis
8	942	52.4	226	2 Q94B43	Q94B43 arabidopsis
9	278	15.5	402	2 Q8S2E6	Q8S2E6 oryza sativ
10	251.5	14.0	750	2 Q851V0	Q851V0 oryza sativ
11	246	13.7	1029	2 Q851V5	Q851V5 oryza sativ
12	239	13.3	337	2 Q8R1D1	Q8R1D1 arabidopsis
13	229	12.8	190	2 Q9ZSH7	Q9ZSH7 arabidopsis
14	229	12.8	190	2 AAQ89626	AAQ89626 arabidops
15	227.5	12.7	461	2 Q9SZA5	Q9SZA5 arabidopsis
16	219.5	12.2	193	2 Q9LJ70	Q9LJ70 arabidopsis
17	215.5	12.0	209	2 Q84R27	Q84R27 arabidopsis
18	214	11.9	306	2 Q851V1	Q851V1 oryza sativ
19	203	11.3	307	2 Q9FUG2	Q9FUG2 arabidopsis
20	197	11.0	519	2 Q851W5	Q851W5 oryza sativ
21	195.5	10.9	301	2 Q9FK61	Q9FK61 arabidopsis
22	192	10.7	334	2 Q9FGD2	Q9FGD2 arabidopsis
23	191	10.6	427	2 Q7X575	Q7X575 oryza sativ
24	182.5	10.2	308	2 Q9LXB1	Q9LXB1 arabidopsis
25	179.5	10.0	302	2 Q9LSP6	Q9LSP6 arabidopsis
26	177.5	9.9	291	2 Q84RE8	Q84RE8 arabidopsis
27	175.5	9.8	545	2 Q851W4	Q851W4 oryza sativ
28	172	9.6	286	2 Q9XIG9	Q9XIG9 arabidopsis
29	172	9.6	389	2 Q9SZ05	Q9SZ05 arabidopsis
30	170.5	9.5	391	2 Q620D2	Q620D2 oryza sativ
31	170.5	9.5	391	2 BAD03253	BAD03253 oryza sat

ALIGNMENTS

RESULT 1

Q8L3W1 PRELIMINARY; PRT; 341 AA.
AC Q8L3W1;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Reduced vernalization response 1.
GN Name=VRN1;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]_
SEQUENCE FROM N.A.
RX MEDLINE=22111275; PubMed=12114624;
RA Levy Y.Y., Mesnage S., Mylne J.S., Gendall A.R., Dean C.;
RT "Multiple roles of Arabidopsis VRN1 in vernalization and flowering time control."
RL Science 297:243-246(2002).
DR EMBL; AF289051; AAM76972.1; -;
DR EMBL; AF289052; AAM76973.1; -;
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 2.
DR PROSITE; PS50863; B3; 2.
SQ SEQUENCE 341 AA; 39274 MW; E40ABFB2C8CC3114 CRC64;

Query Match 100.0%; Score 1796; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 2.1e-129;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPRPFHKLIFSSTIQEKLRLVPDKFKPDKLSVAVALTPDGHVVRVGLRKADNKIW 60
DB 1 MPRPFHKLIFSSTIQEKLRLVPDKFKPDKLSVAVALTPDGHVVRVGLRKADNKIW 60
QY 61 FQDQGEFVDYRISIRIGYLLIFRYEGNSAFSVYIFNLHSEINYHSTGLMDSAHNHFKRA 120
DB 61 FQDQGEFVDYRISIRIGYLLIFRYEGNSAFSVYIFNLHSEINYHSTGLMDSAHNHFKRA 120
QY 121 RLFEDEDEDAEVIIPSSVYPSPLPESTVPANKGYASSAIQTLFTGPVKAEEPTTPKIP 180
DB 121 RLFEDEDEDAEVIIPSSVYPSPLPESTVPANKGYASSAIQTLFTGPVKAEEPTTPKIP 180
QY 181 KRGGRKKNADPEEINNSAPRDDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240
DB 181 KRGGRKKNADPEEINNSAPRDDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240
QY 241 NPFFRVTLRPSLYRGCMYLPSPGPAEKVLSGIFIKVQLAEKQMPVRCLYKAGRAKFS 300
DB 241 NPFFRVTLRPSLYRGCMYLPSPGPAEKVLSGIFIKVQLAEKQMPVRCLYKAGRAKFS 300

```

QY 301 QGWEFTLENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 341
DB 301 QGWEFTLENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 341

RESULT 2
Q6R2U8 PRELIMINARY; PRT; 329 AA.
AC Q6R2U8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Reduced vernalization response 1.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP SEQUENCE FROM N.A.
RA Kwon S.-J., Park B.-S., Kim S.-Y., Choi H.-S., Lee M.-C., Kim J.-S.,
RA Lee S.-I., Lim K.-B., Kim J.-A., Lee M.-R., Jin Y.-M., Kim D.,
RA Kim H.-I.;
RT "Brassica rapa Kwonsim VRN1 mRNA.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY517929; AAR92199.1; -.
SQ SEQUENCE 329 AA; 38058 MW; A24ADB91433459BB CRC64;

Query Match 84.2%; Score 1513; DB 2; Length 329;
Best Local Similarity 85.1%; Pred. No. 1e-107;
Matches 296; Conservative 13; Mismatches 13; Indels 26; Gaps 5;

QY 1 MPRPFHKLIFSSSTIQEKLRLVPDKFVSKFDELSSAVALTVPDGHVWRVGLRKAD--NH 58
DB 1 MPRPFHKLIFSSSTIQEKLRLVPDKFVSKFDELSSAVALTVPDGHVWRVGLRKADNNK 60
QY 59 IFQDGMQEFVDYRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINHYSTGLMDSA--HNH 116
DB 61 IFQDGMQEFVDYRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINHYSTGLMDSASHNNH 120
QY 117 FKARLFEDELEDAEVIFFSSVYPSPLPSTVPANKGYASSAIQTUFTGPVKAEE--PT 174
DB 121 FKARLFEDELEDAETLHTT-----ASAIQSFTGPKPEEATPT 161
QY 175 PTPKPKKGRKKKNAD-PEEINSSAPRDDDPENRSKFYESASARKTVAERERAINA 233
DB 162 QTSKVPKKGKKNADHPPEVNSSAPRDDDPESRSKFYESASARKTVAERERAVNA 221
QY 234 AKTPEPTNPFFRVLRPSLYRGCMYLPSPGFAEKYLSGISGFIKVLAEKQWPVRCLYK 293
DB 222 AKTPEPTNPFFRVLRPSLYRGCMYLPSPGFAEKYLSGISGFIKVLAEKQWPVRCLYK 281
QY 294 AGRAKFSQGWYEFTENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 341
DB 282 AGRAKFSQGWYEFTENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 329

RESULT 4
Q6V0J6 PRELIMINARY; PRT; 329 AA.
AC Q6V0J6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Reduced vernalization response 1.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP SEQUENCE FROM N.A.
RA Kwon S.-J., Park B.-S., Kim S.-Y., Choi H.-S., Lee M.-C., Kim J.-S.,
RA Lee S.-I., Lim K.-B., Kim J.-A., Hong K.-Y., Lee M.-R., Jin Y.-M.,
RA Kim D., Kim H.-I.;
RT Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY356368; AAQ55453.1; -.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 2.
DR PROSITE; PS50863; B3; 2.
SQ SEQUENCE 329 AA; 37931 MW; BFF71D55ADE2655A CRC64;

Query Match 84.0%; Score 1509; DB 2; Length 329;
Best Local Similarity 85.1%; Pred. No. 2e-107;
Matches 296; Conservative 12; Mismatches 14; Indels 26; Gaps 5;

QY 1 MPRPFHKLIFSSSTIQEKLRLVPDKFVSKFDELSSAVALTVPDGHVWRVGLRKAD--NH 58
DB 1 MPRPFHKLIFSSSTIQEKLRLVPDKFVSKFDELSSAVALTVPDGHVWRVGLRKADNNK 60
QY 59 IFQDGMQEFVDYRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINHYSTGLMDSA--HNH 116
DB 61 IFQDGMQEFVDYRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINHYSTGLMDSASHNNH 120
QY 117 FKARLFEDELEDAEVIFFSSVYPSPLPSTVPANKGYASSAIQTUFTGPVKAEE--PT 174
DB 121 FKARLFEDELEDAETLHTT-----ASAIQSFTGPKPEEATPT 161
QY 175 PTPKPKKGRKKKNAD-PEEINSSAPRDDDPENRSKFYESASARKTVAERERAINA 233
DB 162 QTSKVPKKGKKNADHPPEVNSSAPRDDDPESRSKFYESASARKTVAERERAVNA 221
QY 234 AKTPEPTNPFFRVLRPSLYRGCMYLPSPGFAEKYLSGISGFIKVLAEKQWPVRCLYK 293
DB 222 AKTPEPTNPFFRVLRPSLYRGCMYLPSPGFAEKYLSGISGFIKVLAEKQWPVRCLYK 281
QY 294 AGRAKFSQGWYEFTENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 341
DB 282 AGRAKFSQGWYEFTENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 329

RESULT 3
AAR92199 PRELIMINARY; PRT; 329 AA.
AC AAR92199;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Reduced vernalization response 1.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]

```

```
Db 61 INFQDQWQFVDRYSIRIGYLLIFRYEGNSAFSVCYNLPQSEINHYSTGLMDSASHNNH 120
QY 117 FKARLFEDLEDEDAEVIFFPSVYPSLPSTVPANKGYASSAIOTLFTGPVKAEB--PT 174
Db 121 FKPRLFEDLEDEDAETLHTT-----ASAIQSFFTGPGVKEAPTPT 161
QY 175 PTPKIPKGRKKKNAD-PEINSSAPRDDDDPENRSKPFYESASARKRTVTAEERERAINA 233
Db 162 QTSKVPKGRKKKNADHPEEVNSSAPRDDDDPENRSKPFYESASARKRTVTAEERERAVNA 221
QY 234 AKTFEFTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSIFIKVQLAEKQWVRCLYK 293
Db 222 AKTFEFTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSIFIKVQLAEKQWVRCLYK 281
QY 294 AGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 341
Db 282 AGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 329

RESULT 5
AAQ55453
ID AAQ55453 PRELIMINARY; PRT; 329 AA.
AC AAQ55453;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE Reduced vernalization response 1.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3711;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Samjin;
RA Kwon S.-J., Park B.-S., Kim S.-Y., Choi H.-S., Lee M.-C., Kim J.-S.,
RA Lee S.-I., Lim K.-B., Kim J.-A., Hong K.-Y., Lee M.-R., Jin Y.-M.,
RA Kim D., Kim H.-I.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY356368; AAQ55453.1; -- BFF71D55ADE2655A CRC64;
SQ SEQUENCE 329 AA; 37931 MW; 37931 MW; 37931 MW; 37931 MW; 37931 MW;

Query Match 84.0%; Score 1509; DB 2; Length 329;
Best Local Similarity 85.1%; Pred. No. 2e-107;
Matches 296; Conservative 12; Mismatches 14; Indels 26; Gaps 5;

QY 1 MPRPFHKLIFSSTIQEKLRLVDPKFSKFDLSVAVALTPDGHVVRVGLRKAD--NK 58
Db 1 MPRPFHKLIFSSTIQEKLRLVDPKFSKFDLSVAVALTPDGHVVRVGLRKADNNK 60
QY 59 INFQDQWQFVDRYSIRIGYLLIFRYEGNSAFSVCYNLPQSEINHYSTGLMDSASHNNH 116
Db 61 INFQDQWQFVDRYSIRIGYLLIFRYEGNSAFSVCYNLPQSEINHYSTGLMDSASHNNH 120
QY 117 FKARLFEDLEDEDAEVIFFPSVYPSLPSTVPANKGYASSAIOTLFTGPVKAEB--PT 174
Db 121 FKPRLFEDLEDEDAETLHTT-----ASAIQSFFTGPGVKEAPTPT 161
QY 175 PTPKIPKGRKKKNAD-PEINSSAPRDDDDPENRSKPFYESASARKRTVTAEERERAINA 233
Db 162 QTSKVPKGRKKKNADHPEEVNSSAPRDDDDPENRSKPFYESASARKRTVTAEERERAVNA 221
QY 234 AKTFEFTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSIFIKVQLAEKQWVRCLYK 293
Db 222 AKTFEFTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSIFIKVQLAEKQWVRCLYK 281
QY 294 AGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 341
Db 282 AGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 329

RESULT 6
```

```
Q9LJ67
ID Q9LJ67 PRELIMINARY; PRT; 230 AA.
AC Q9LJ67;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Gb|AAD43153.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones."
RL DNA Res. 7:217-221(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP000735; BAB01695.1; --
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 1.
DR PROSITE; PS50863; B3; 1.
SQ SEQUENCE 230 AA; 26219 MW; B96A826B384C9F3C CRC64;

Query Match 66.6%; Score 1197; DB 2; Length 230;
Best Local Similarity 99.1%; Pred. No. 1e-83;
Matches 230; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 110 MDSAHNHFKARLFEDLEDEDAEVIFFPSVYPSLPSTVPANKGYASSAIOTLFTGPVK 169
Db 1 MDSAHNHFKARLFEDLEDEDAEVIFFPSVYPSLPSTVPANKGYASSAIOTLFTGPVK 60
QY 170 ABEPTPTPKIPKGRKKKNADPEINSSAPRDDDDPENRSKPFYESASARKRTVTAEERER 229
Db 61 --EPTPTPKIPKGRKKKNADPEINSSAPRDDDDPENRSKPFYESASARKRTVTAEERER 118
QY 230 AINAAKTEPTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSIFIKVQLAEKQWVR 289
Db 119 AINAAKTEPTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSIFIKVQLAEKQWVR 178
QY 290 CLYKAGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 341
Db 179 CLYKAGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRFVLKVTAFRVNEYV 230

RESULT 7
Q9XIB5
ID Q9XIB5 PRELIMINARY; PRT; 226 AA.
AC Q9XIB5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F13F21.8 protein.
GN Name=F13F21.8;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federici N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luo S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
```



```
QY 197 SSAPRDDPNRSKFSYASARKTVAERERAINAAKTFTPTNPFVRLRPSVLYRG 256
DB 248 ASIP-----PSIRA--YKGVSRRAVATAERQATEIAHAFRSPLYCVRIMSTMVYYS 301
QY 257 CMIYLPSPGAEXYL-SGISGFIKVLAEKQWVRCLYKAG-RAKFSQGWYFTLENNLGE 314
DB 302 FMRPFTGSRQHLPRRTDVLVRDGGKWSV--LYPNTRDLRSGMCAFAARGNCLSE 359
QY 315 GDVCPVELLRTRDFVLKVTAFRVNE 339
DB 360 GDVCPVELVAARF--RVHIFRVVE 382

RESULT 10
Q851V0 PRELIMINARY; PRT; 750 AA.
AC Q851V0;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Putative auxin response factor.
GN Name=OSJNB0111B07.22;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Ziemann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.O., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC097280; AAO34502.1; -
DR Gramene; Q851V0; -
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 4.
DR PROSITE; PS50863; B3; 4.
SQ SEQUENCE 750 AA; 83502 MW; E9E3980EF8EF3E3 CRC64;

Query Match 14.0%; Score 251.5; DB 2; Length 750;
Best Local Similarity 26.9%; Pred. No. 1.5e-10;
Matches 86; Conservative 55; Mismatches 134; Indels 45; Gaps 13;

QY 22 VPDKFSVKFDLSVAVALTPDGHVVRVGLKAD-NK1FQDQWQEFVDRYSIRIGYLL 80
DB 424 VPARFANNFGHISEEVNLRSPSGETWSTGVANSAGELVLPQWKEFVGDNGISEGDC 483
QY 81 IFRYEG-NSAFSVYIFNLSSHSEINYSHTGLMSAHNHFKRRLPEDEDE-DAEVIFFSS 138
DB 484 LFRYSVSSSFDVLPDPSCGE-----KASPHFVSGHFGRAENSAGAEOGGRNG 533
QY 139 VYPSPESTVPANKGYASSAQTFLTPGVKAAEPTPTPKIPK--KRG-----RKKGN 189
DB 534 RTTPPI-----VDGNGRHHLEWTLHRNSCRS-----IPACKRSIFSDTEAKEND 581
QY 190 ADEEINSGAPRDDDENRSKYESASAR--KRTVTAERERAINAAKTFTPTNPFV 247
DB 582 GEDEDVAAA-----EGGRYGEYFSGHGRVAEYNLRREDEBISRVVPVQPCNPFVQV 637
QY 248 LPSVYL--YRGCMYLPSCFAKYLSGISGFIKVLAEK--QWVRCLYKAG-RAKFSQ 302
DB 638 IHSSHVRSSKYCIVGVSPEFAGKYLGAVERVVLRAERSGGEMHVPFVHRQNTRGFYAG 697
```

```
QY 303 WYEFTLENNLGEQDVCVFEL 322
DB 698 WRQFAGDNLRLVAHDVCLPEL 717

RESULT 11
Q851V5 PRELIMINARY; PRT; 1029 AA.
AC Q851V5;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Putative auxin response factor.
GN Name=OSJNB0111B07.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Ziemann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.O., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC097280; AAO34491.1; -
DR Gramene; Q851V5; -
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 5.
DR PROSITE; PS50863; B3; 5.
SQ SEQUENCE 1029 AA; 118189 MW; 206B6AD0CA4B7645 CRC64;

Query Match 13.7%; Score 246; DB 2; Length 1029;
Best Local Similarity 23.8%; Pred. No. 5.9e-10;
Matches 74; Conservative 60; Mismatches 123; Indels 54; Gaps 8;

QY 5 PFHKLIFSSTIQEKLVRVPDKFVSFKDLSVAVALTPDGHVVRVGLRKADNKIWFQDS 64
DB 149 YFFKVMIGGF--RRQMTIPYFAENFRDQIQGTIKLKARNGTCSVLVDKCSNKLVLTKG 206
QY 65 WQEFVDYSIRIGYLLIFRYEGNSAFSVYIFNLSSHSEINYSHTGLMSA-HNHFKRRLF 123
DB 207 WAEFANSHDIXMGDFLVRFYTGNSQFEVKIFDPS-----GCVKAASHNAVNIQHA 257
QY 124 EDLEDEDAEVIFFPSVYPSPIPESTVPANKGYASSAQTFLTPGVKAAEPTPTPKIPK 183
DB 258 QNMQGDPIEILSCSDEHLRAQSLTTERQNQ-----PEKD 291
QY 184 GRKKGNADPEEINSGAPRDDDENRSKYESASARKTVAERERAINAAKTFTPTNPF 243
DB 292 VIDNCNCKKKKTEHASSSEDQ-----ETPTAEVHRMKVEEMVRAIHS-----NHVP 337
QY 244 FRVLRLPSVLYR-GCMYLPSCFAKYLSGISGFIKVLAEKQWVPR-CLYKAGRAKFSQ 301
DB 338 FVAVVKKSNVTPQPCYVAISRKYANEYFPGDQMLTLQRHGKRWQVKFCISKKLEMLSK 397
QY 302 GWYFTLENNL 312
DB 398 GWRKFTRDNEL 408

RESULT 12
Q8RYD1 PRELIMINARY; PRT; 337 AA.
ID Q8RYD1
```


Query Match 12.8%; Score 229; DB 2; Length 190;
Best Local Similarity 33.3%; Pred. No. 1.2e-09;
Matches 53; Conservative 25; Mismatches 51; Indels 30; Gaps 4;
QY 6 FHKLISSIOEKRLAVPDKFVSKFKDELSVAVALTPDGHVVRVGLRKADNKNIFQDGH 65
DB 30 FFKLVLPSTWKOQOMRIPFRFVKLOGSKLSEVTVLTPAGYKRSIKLKRIGEIEWFHEG 89
QY 66 QEFVDYSYRIGYLLIFRYEGNSAFSVYFNLSHSEINY-----HSTGL- 109
DB 90 SEFAEASHSEEGHFLFEYKNSFRVILFNASACETNPLDAVHIDSDDDVIEITGKE 149
QY 110 MOSAHNFKRAR-----LFEDL-----EDED AEVI 134
DB 150 FDEHKSKRPRDIEFDKILHDVDVMQVLKEEEDKRVL 188

RESULT 15

Q9SZA5 PRELIMINARY; PRT; 461 AA.
AC Q9SZA5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein F17M5.40 (Hypothetical protein AT4g33280).
GN Names=F17M5.40; Synonyms=AT4g33280;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Rose M., Hempel S., Entian K.-D., Hoheisel J., Mewes H.W.,
RA Mannhaupt G., Mayer K.F.X., Schueller C.;
RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035678; CAB38786.1; -;
DR EMBL; AL161583; CAB80045.1; -;
DR PIR; T05979; T05979.
DR GO; GO:0003677; F.DNA binding; IEA.
DR GO; GO:0006355; P.regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 2.
DR PROSITE; PS00863; B3; 2.
KW Hypothetical protein.
SQ SEQUENCE 461 AA, 52997 MW, D7EB497600A33062 CRC64;

Query Match 12.7%; Score 227.5; DB 2; Length 461;
Best Local Similarity 23.3%; Pred. No. 5.3e-09;
Matches 86; Conservative 63; Mismatches 139; Indels 81; Gaps 13;
QY 22 VPDKFKSKDELSVAVALTPDGHVVRVGLRKADNK-IWFQDQGEFVDYSIRIGYLL 80
DB 108 IPRKFTCKRKLQVITLSPSGVTYNGVEEDEDKTAFRGWDKFKVDHSLSEENDLL 167
QY 81 IFRYEGNSAFSVYFN-----LSHSE-----INHSTGLMDSADNHFKRA 120
DB 168 VFKFGVSEFEVLVPGQILCEKPTSYFVRKCGHAETKGIIDFNATS-SRSPKRFHND 226
QY 121 RL-----FEDLEDAEVIFFSPVPSPLPESTVPANKGYASSAIQTL 163

Search completed: December 30, 2004, 06:10:40
Job time : 100 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 19:00:56 ; Search time 23 Seconds
(without alignments)
1426.518 Million cell updates/sec

Title: US-10-088-187A-11
Perfect score: 1796
Sequence: 1 MRPFPFKLIFSSTIQKRL.....LLTRDFVLKVTAFRVNEYV 341
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	946	52.7	226	E96531	hypothetical prote
2	229	12.8	190	T02015	hypothetical prote
3	227.5	12.7	461	T05979	hypothetical prote
4	182.5	10.2	308	T49920	hypothetical prote
5	172	9.6	286	H84537	hypothetical prote
6	172	9.6	389	T04786	hypothetical prote
7	161.5	9.0	352	T05106	hypothetical prote
8	145	8.1	344	T05104	hypothetical prote
9	141.5	7.9	210	D96531	hypothetical prote
10	141	7.9	283	T45871	hypothetical prote
11	140.5	7.8	461	T05101	hypothetical prote
12	137.5	7.7	490	T01545	hypothetical prote
13	134	7.5	512	T05107	hypothetical prote
14	133.5	7.4	984	G86393	protein T24P13.6 [
15	133.5	7.4	1440	G86393	hypothetical prote
16	133	7.4	899	H84639	hypothetical prote
17	132	7.3	493	T05105	hypothetical prote
18	118	6.6	478	T05102	hypothetical prote
19	116.5	6.5	851	F84639	hypothetical prote
20	111	6.2	608	AE2531	hypothetical prote
21	110.5	6.2	1021	T05108	hypothetical prote
22	107.5	6.0	2783	A41948	alpha-fetoprotein
23	104	5.8	497	T14433	reproductive meris
24	102.5	5.7	134	E86301	F19K19.7 protein -
25	102.5	5.7	232	T22698	hypothetical prote
26	101.5	5.7	243	TJ0788	tonB protein - Kle
27	101.5	5.7	482	T22981	hypothetical prote
28	101	5.6	1507	T18544	alpha-2-macroglobu
29	99	5.5	298	T37251	homeobox protein c

30	98	5.5	586	2	T29657	hypothetical prote
31	98	5.5	720	2	T36819	hypothetical prote
32	97.5	5.4	1082	2	T15269	hypothetical prote
33	97	5.4	252	2	A00267	tonB protein [impo
34	97	5.4	525	2	T05109	hypothetical prote
35	96.5	5.4	297	2	A84767	hypothetical prote
36	96	5.3	852	2	T06310	hypothetical prote
37	95.5	5.3	616	2	G96511	hypothetical prote
38	95	5.3	673	1	S73444	MG032 homolog B01
39	95	5.3	932	2	T45894	hypothetical prote
40	94.5	5.3	1003	2	T16740	hypothetical prote
41	94	5.2	1008	2	S38003	translation elonga
42	92.5	5.2	599	2	T00948	hypothetical prote
43	92	5.1	541	2	C81357	probable secreted
44	92	5.1	682	2	G84639	hypothetical prote
45	92	5.1	823	2	T09882	heat shock protein

ALIGNMENTS

RESULT 1

E96531

hypothetical protein F13F21.8 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: E96531

R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96531

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-226 <STO>

A:Cross-references: UNIPROT:Q9XIB5; GB:AE005173; NID:g5430753; PIDN:AD43153.1; GSPDB:G

C:Genetics:

A:Gene: F13F21.8

A:Map position: 1

Query Match 52.7%; Score 946; DB 2; Length 226;
Best Local Similarity 81.2%; Pred. No. 3.3e-69;
Matches 190; Conservative 14; Mismatches 18; Indels 12; Gaps 5;

QY 110 MDSAHNF-KEARLPEDELEDAEVIFFSSVYPSLPSTVPANKGY-ASSAIQILFTCP 167

Db 3 MDSAQNQFNKARLPEDPELDAKVIYPSN-----PESTFVNKGYGGSTAIQSFFK-E 55

QY 168 VKAEPTPTPKPKKGRKKKNADPEEINSSAPRDDDDPENRSKFYESASARKTSTAER 227

Db 56 SKAEE---TPVKLKGRRKKKNPNEEVNNSSTPGDDSENRSKFYESASARKTSTAER 112

QY 228 ERANNAKTFFTPNPFVRLRPSLYRGCMYLPSPGFAEKYLSGIFGFKVQLAEKQWP 287

Db 113 ERAVNAKTFFTPNPFVRLRPSLYRGCMYLPSPGFAEKYLSGIFGFKVQLAEKQWP 172

QY 288 VRCLYKAGRAKPSQGWYEFTEENNIGEGDVCVFELLRTDRFVLKVTAFRVNEYV 341

Db 173 VRCLYKAGRAKPSQGWYEFTEENNIGEGDVCVFELLRTDRFVLKVTAFRVNEYV 226

RESULT 2

T02015

hypothetical protein T15B16.18 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004

C;Accession: T02015
R;Stonking, T.; Smith, R.
submitted to the EMBL Data Library, November 1998
A;Description: The sequence of A. thaliana T15B16.
A;Reference number: Z14488
A;Accession: T02015
A;Species: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-190 <STO>
A;Cross-references: UNIPROT:Q9ZSH7; EMBL:AF104919; NID:g3859590; PID:g3859591
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 44/3; 171/1
A;Note: T15B16.18

Query Match 12.8%; Score 229; DB 2; Length 190;
Best Local Similarity 33.3%; Pred. No. 2.8e-11;
Matches 53; Conservative 25; Mismatches 51; Indels 30; Gaps 4;

QY 6 FFKLIFSSTIOEKRLRVDPKFSKDELVAVALTPDGHVVRVGLRKADNKIWFQDQW 65
DB 30 FFKLVPTMKDKXMRIPPRFVKLGSKLSEVTVITPAGYKRSIKLKRIGEEIWFHEGW 89

QY 66 QEFVDYRYSIRIGYLIFRYEGNSAFSVYFNLHSEINY-----HSTGL- 109
DB 90 SEFAEAHSIEGHFLFPEYKXNSPRVIFNACETNYPLDAVHIIDSDDDVDIEITGKE 149

QY 110 MDSAHNHFKRAR-----LFDL-----EDEDAEVI 134
DB 150 FDTHEKSKRPRDIEFDKILHDVDVMQVLKDEEDKRVL 188

RESULT 3
T05979
hypothetical protein F17M5.40 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05979
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hoheisel, J.; Mewes, H.W.; Mayer, K.P.X.
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15263
A;Accession: T05979
A;Molecule type: DNA
A;Residues: 1-461 <BEV>
A;Cross-references: UNIPROT:Q9SZA5; EMBL:AL035678; GSPDB:GN00062; ATSP:F17M5.40
A;Experimental source: cultivar Columbia; BAC clone F17M5
C;Genetics:
A;Gene: ATSP:F17M5.40
A;Map position: 4
A;Introns: 37/3; 328/1; 362/3

Query Match 12.7%; Score 227.5; DB 2; Length 461;
Best Local Similarity 23.3%; Pred. No. 1.2e-10;
Matches 86; Conservative 63; Mismatches 139; Indels 81; Gaps 13;

QY 22 VPDKVSFKDELVAVALTPDGHVVRVGLRKADNKIWFQDQWQEFVDYRYSIRIGYL 80
DB 108 IPRKESTCKRKLQIVTLKSPSGVTVYNGVEEDKTMARFGWDKFKDHSLENDLL 167

QY 81 IFRYEGNSAFSVYIFN-----LSHSE-----INYHSTGLMDSAHNHFKRA 120
DB 168 VFKHGVSEFVLVDGDTLCEKPTSYFVRKCGHAETKGIIDFNATS-SRSPKRHFND 226

QY 121 RL-----FEDLEDAEVIFFSPSVPSLPPESTVPANKGYASSAIOTL 163
DB 227 DVETTPNOOLVISPVDNELEDLIDFLDKTLNLLVASHFGYQSEHNSDIDTA 286

QY 164 FTGPVKAEEPTTPKIP-----KKRGRKKKNADPEEINSSAPRDDDPENRSKFYS 214
DB 287 -SAQLPVSIPSTVRSVGKYPKPLSGFKKWRSLNDNDLQ-----KA 327

QY 215 ASARKRTVTAERERAINAAKTFETPNPFVRVLRPSLYRGCTIMYLPSPGFAEKLSGIS 274

C;Accession: T02015
R;Stonking, T.; Smith, R.
submitted to the EMBL Data Library, November 1998
A;Description: The sequence of A. thaliana T15B16.
A;Reference number: Z14488
A;Accession: T02015
A;Species: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-190 <STO>
A;Cross-references: UNIPROT:Q9ZSH7; EMBL:AF104919; NID:g3859590; PID:g3859591
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 44/3; 171/1
A;Note: T15B16.18

Query Match 10.2%; Score 182.5; DB 2; Length 308;
Best Local Similarity 24.6%; Pred. No. 3.1e-07;
Matches 83; Conservative 51; Mismatches 145; Indels 59; Gaps 13;

QY 5 FFKHLIFSSTIOEKRLRVDPKFSKFD-ELSAVALTPDGHVVRVGLRKADNKIWFQD 63
DB 19 FFKVLKSDMSSEDTAIPYDFVINFSDELSDGMKFRVQNGSNKVKISK-NPRFYFME 77

QY 64 --GWOEFDYRYSIRIGYLIFRYEGNSAFSVYFNLHSEINYHSTGLMDSAHNHFKRA 121
DB 78 KSGWEKFIIDNALGDHFEFLTFTHGQMSFTVKIFNKGDKEMWPPQSPASPASS--SRVK 135

QY 122 LFEDLEDAEVIFFSPSVPSLPPESTVPANKGYASSAIOTLFTGPVKAEEPTTPKIPK 181
DB 136 TEQDVKRE--EEVLVSSDSRSRGPTTAETNRG--GSYKRLNFGKKAEB-TQYK--- 187

QY 182 KRGKKKNADPEEINSSAPRDDDPENRSKFYSASARKRTVTAERERAINAAKTF--EP 239
DB 188 -----RTERTQNSKRTVRSVSKERVYAGEP 212

QY 240 TNPF--FRVLRPSLYRGCTIMYLPSPGFAEKLSGISGFIKVL--AEKQWPVRLYKAG 295
DB 213 SSSVAGFKIFTSKYSIKS---LAIPKPG-NYMPKETRVRKIHHPDGEKTKWVFWVKER 268

QY 296 RAKFSQGYEFTLNNLGEQDVCVFELLRTDFVLKVT 333
DB 269 GQIFSGGKRLCKEYVPVFGDTCTLTITPLELLVWS 306

RESULT 5
H04537
hypothetical protein At2g16210 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H04537
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: H84537

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-286 <STO>

A:Cross-references: UNIPROT:Q9XIG9; GB:AE002093; NID:g4678216; PIDN:AAD26962.1; GSPDB:GN

C:Genetics:

A:Gene: At2g16210

A:Map position: 2

C:Superfamily: Arabidopsis thaliana hypothetical protein F17114.30

Query Match 9.6%; Score 172; DB 2; Length 286;
Best Local Similarity 21.3%; Pred. No. 2e-06;
Matches 72; Conservative 56; Mismatches 128; Indels 82; Gaps 13;

QY 6 FHLKIFSSIT-DEKRLRVPDKFVSKFD-ELSVAVALTVPDGHVVRVGLRKADNKIWFOD 63

DB 19 FFAVQSVINVSSENKALPHDSRSTDELGRKMKIRAQWNGSMEVGIK-NPRIFYME 77

QY 64 --GWQEFVDYRISIRIGYLLIFRYEGNSAFSVYIFNLHSEINYNHSTGLMDSAHNHFKRAR 121

DB 78 KSGWEKFRVDNALGNSSELLTFTHKGWHTVNIPLKDGKE-----NMQPPQSRSPFFAS 130

QY 122 LPEDLEDEDAEVIFPSSVPSPLPESTVPANKYASSAIQTLTGPVKAEEPTPTPKIPK 181

DB 131 SKQBEENDIKERVV-----VSSNRGQTAA-----E 155

QY 182 KGRKKKNADPEINSSAPRDDDDPENRSKFYSASARKETVTAEEERAINAAKTF---- 237

DB 156 SKGRKL-----NLGK-----RAKESQSKTKTEKVVRRARSVDAGAS 191

QY 238 BPTNPFVRVLRPSVLYRGIMYLPSPGFAEKYLGSIGFIKQV--LAERQWVPVRCILYKAG 295

DB 192 SSTAAFTILFQGYL--VFLRIPNSVSKQVPDEKTVFKIHHPNGKSNVVLIRFG 248

QY 296 RAKFSQGWTEFTLENNLGGDVCVFELLTRDFVLKVT 333

DB 249 --AFSGWRVRVKEVPLAVGDTCKFTFKPKELLVVS 284

RESULT 6

T04786

hypothetical protein F10M10.170 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: T04786

R:Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Hoheisel, J.; Mewes, H.W.; Mayer,

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15384

A:Accession: T04786

A:Molecule type: DNA

A:Residues: 1-389 <BEV>

A:Cross-references: UNIPROT:Q9SZ05; EMBL:AL035521

A:Experimental source: cultivar Columbia; BAC clone F10M10

C:Genetics:

A:Map position: 4

A:Introns: 29/3; 154/1; 203/1; 276/1; 306/3

A:Note: F10M10.170

Query Match 9.6%; Score 172; DB 2; Length 389;
Best Local Similarity 21.0%; Pred. No. 3.1e-06;
Matches 86; Conservative 51; Mismatches 154; Indels 118; Gaps 16;

QY 1 MRPFPFKLIPSTIQEKLRLVPDKFVSKFDELSVAVALTVPDGHVVRVGLRKADNKI 60

DB 12 LPPRFF--TVFVSHFSSFMWIPVSVYDIHPRFPKTVILRGPGGCKWVATEIKDDEVL 69

QY 61 FQDQWQEFVDYRISIRIGYLLIFRYEGNSAFSVYIF-----N 96

DB 70 FQGGHPKFRVDNLNDGDFLTAYNGAHIFVSIIRFGYDACKIEISVTELEBEEDSVIS 129

QY 97 LSHSEINYNHSTGLMDSAHNHFKRAR-----LPEDLEDEDAEVIFPSSVPSPLPEST 148

DB 130 LSESDT---TGAKSEMKNVTPEGRDKGSKVVEVEDSDDDDEE-----DSVYSESEETE 182

QY 149 VPANKGYASSAIQTLTGPVKAEEPTPTPKIPK--KGRKK----- 187

DB 183 TDTSEF-----KVAKT-----IPKQKGGKKKQVQVSSDEDEBEDSDS 225

QY 188 -----KNADPEINSSAPRDDDDPENRS-----KFYSASARKTV--TAEERER 229

DB 226 DVIETFGQLDIBENGISEEDSYAPDKEDTATASFVKPKVANKVANLKKKVDPMKNP 285

QY 230 AINAAKT--FEPT--NPFRVVRVLRPSVLYRGIMYLPSPGFAEKYLGSIGFIKVLAEKQ 285

DB 286 YLDDPKNIHFETNVKNRLYELVHAQLVKDYCLRF-----GDYNYIDRFGLKSAKTAK 339

QY 286 WVR--CLYKAGRAKFSQGWTEFTLENNLGGDVCVFELLTRDFVLKV 332

DB 340 WKDQRCVTKR-----WNRICKRNKLKEDRILCELLRKGTFVYAI 379

RESULT 7

T05106

hypothetical protein F28M20.170 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: T05106

R:Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Hoheisel, J.; Mewes,

submitted to the Protein Sequence Database, November 1998

A:Reference number: Z15398

A:Accession: T05106

A:Molecule type: DNA

A:Residues: 1-352 <BEV>

A:Cross-references: UNIPROT:Q9SB80; EMBL:AL031004

A:Experimental source: cultivar Columbia; BAC clone F28M20

C:Genetics:

A:Map position: 4

A:Introns: 26/3; 125/1; 166/3

A:Note: F28M20.170

C:Superfamily: Arabidopsis thaliana hypothetical protein F28M20.180

Query Match 9.0%; Score 161.5; DB 2; Length 352;
Best Local Similarity 20.1%; Pred. No. 1.9e-05;
Matches 69; Conservative 49; Mismatches 102; Indels 123; Gaps 15;

QY 5 FFHKLII--FSSITQEKRLRVPDKFVSKF---KDELSVAVALTVPDGHVVRVGLRKADNKI 59

DB 14 FQPTILPGFKS-----HIKIPVKFSTKIEGHEGNTVLRSPSPRTWKV--KMEGHKL 66

QY 60 WPDQWQEFVDYRISIRIGYLLIFRYEGNSAFSVYIFNLHSEINYNHSTGLMDSAHNHFKR 119

DB 67 --TEGWEKFEVAHDLRVGDFVVFVKHGDMLFHVTAIGPSCCEVQY----- 109

QY 120 ARLFEDLEDEDAEVIFPSSVPSPLPESTVPANKGYASSAIQTLTGPVKAEEPTPTPKI 179

DB 110 -----APSRSHD----- 116

QY 180 PKKGRKKKNADPEINSSAPRDDDDPENRSKFYSASARKRTVTAEERERAINAAKTFEP 239

DB 117 -----RNESEDEIGSS-----RNEKILIE-----ENKTEPDQFSP 147

QY 240 TNPFRVVRVLRPSVLYRGIMYLPSPGFAEKYLGSIGS-GFIKVLAEKQ-----WVPRC-LYKA 294

DB 148 DLTCFSQSQTASNLTRD-LVGIPRDFAKRY--GLNIGRHEIVLMDSEGNTSESEVSKYS 204

QY 295 GRAKFSQGWTEFTLENNLGGDVCVFELLTRDFVLKVTAFRV 337

DB 205 GRVFIAGGWTSLCTANKLEVGDSCFTKLLRN-----KIPVPRL 243

RESULT 8

T05104

hypothetical protein F28M20.150 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

submitted to the Protein Sequence Database, November 1998
A:Reference number: Z15398

A:Accession: T05101

A:Molecule type: DNA

A:Residues: 1-461 <BEV>

A:Cross-references: UNIPROT:O81778; EMBL:AL031004

A:Experimental source: cultivar Columbia; BAC clone F28M20

C:Genetics:

A:Map position: 4

A:Introns: 26/3; 125/1; 163/3; 245/3; 316/3

A:Note: F28M20.120

C:Superfamily: Arabidopsis thaliana hypothetical protein F28M20.180

Query Match

Best Local Similarity 7.8%; Score 140.5; DB 2; Length 461;

Matches 81; Conservative 50; Mismatches 151; Indels 95; Gaps 18;

QY 5 FFKLI--FSSITQKRLVDPKFSKF---KDELSVALVTPDGHVVRV---GLRKAD 56

DB 14 FFOPLLPGFQS-----NLKIPVNYFSEHTEGHEGKTVTLRTDASERTWEVRMEGHR--- 65

QY 57 NKIFQDQWQFVDYRSIRIGVLLIFRYEGNSAFSVIFNLGSHSEINY-----HSTGLMD 111

DB 66 ----LTEGKWEFVEADHLRIGDFVFRHGEDMVHVTALGPSCCHIQYQSSRHEEG-EE 120

QY 112 SAHNHFKRARLFEDLEDAEV-----IFPSSVYPSLPPESTVPANKGYASSAIQTFTG 166

DB 121 SCENEISEKEGENVQKESDKSSDLNCFPSQSVTHSNISRDVAVPRDFV----- 170

QY 167 PVKAEPTTPKIPKGRKKGNADPEEINSSAPRDDDDPENRSKFYESAS---ARKRTVT 223

DB 171 -----KRSFGSKGR---HEIVLMEEGKSWESVSKYMSGAVYLVGWTTF 213

QY 224 AEERERAINAAKTF---BPTNPFPRVLRVPSVLYRGCIWYLPSPGFAEKYLSIGSFI-- 277

DB 214 CTENKLDVGDSTFKLLQAKTPVQLCSRTK-----HPLSFTK-----VNGLINP 260

QY 278 -KVLAEK---QWPVRLCYKAGRAKFSQSG---WYEFTLENNLGEQDVCVFELLR----- 324

DB 261 GKILLVDKDAEWSMMLKVDSRGAVYIIGNDWKSCFAANEVGAESLALELIQGVLLN 320

QY 325 ---TRDFVLKVTAFRNE 339

DB 321 QITTCFQWQPSFKAE 337

RESULT 12

T01545

hypothetical protein A_IG005110.15 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004

C:Accession: T01545

R:Andrews, S.

submitted to the EMBL Data Library, July 1997

A:Description: The sequence of A. thaliana IG005110.

A:Reference number: Z14347

A:Accession: T01545

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-490 <AND>

A:Cross-references: UNIPROT:O23076; EMBL:AF013293; NID:g2252823; PID:g2252835

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Introns: 112/3; 216/1; 256/3; 391/3

A:Note: A_IG005110.15

C:Superfamily: Arabidopsis thaliana hypothetical protein F28M20.180

Query Match

Best Local Similarity 7.7%; Score 137.5; DB 2; Length 490;

Matches 81; Conservative 42; Mismatches 158; Indels 65; Gaps 16;

QY 22 VPDKFVSKF---KDELSVALVTPDGHVVRVGLRKADNKIWFQDQWQFVDYRSIRIGY 78

DB 6 IFVAFFLKYIEGRYEQKTAKLASDASKRTWEY---KIDQG-RLTDGKWEFAVSHDLRIGD 61

QY 79 LLIFRYEGNSAFSVIFNLGSHSEINYHSTGLM-----DSAHNHFKRARLFE 124

DB 62 IVVFRQESDLAFHVTLGLGPGSCGCIQYGSCSVKKNLGDGELCSVSGSSRNVRFPFRGFVRD 121

QY 125 DLEDEDAEVI-----FPSSVYPSLPPESTVPANKGYASSAIQTFTGVPKAESEPT 174

DB 122 NGVVGSGEIVLMNEKGRSNFNLKQPS---NGTVVRGGMVSFC---DANGLKAGD-N 173

QY 175 PTPKIPKGRG---RKKNADPEEINS---SAPRDDDDPENRSKFYESASAKRTVTABE 226

DB 174 YTFKLKIRAGTLVLRLLPNPEKEANEVSLPEPESDAERNLEKIQKKEV-KKNVTRE- 231

QY 227 RERAINAAKTFPTNPFPRVLRVPSVLYRGCIWYLPSPGFAEKYLSIGI-----SGFIKVQLA 282

DB 232 -----AESQDPSCFVANVSPSSL-RYDTLYLPKFRMRB--NGVDKRCGEMILINEK 281

QY 283 EKQWVPRCLYK--AGRAKFSQGYEFTLENNLGEQDVCVFELLRTR 326

DB 282 GKSMTLDLKVKSCTSLIKGRWSFCSANGLRAGSIITLKLKKR 327

RESULT 13

T05107

hypothetical protein F28M20.180 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: T05107

R:Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Hoheisel, J.; Mewes,

submitted to the Protein Sequence Database, November 1998

A:Reference number: Z15398

A:Accession: T05107

A:Molecule type: DNA

A:Residues: 1-512 <BEV>

A:Cross-references: UNIPROT:Q9SB79; EMBL:AL031004

A:Experimental source: cultivar Columbia; BAC clone F28M20

C:Genetics:

A:Map position: 4

A:Introns: 26/3; 184/3; 322/3; 402/3

A:Note: F28M20.180

C:Superfamily: Arabidopsis thaliana hypothetical protein F28M20.180

Query Match

Best Local Similarity 7.5%; Score 134; DB 2; Length 512;

Matches 69; Conservative 47; Mismatches 111; Indels 120; Gaps 12;

QY 1 MPRP---FFHKLIFSSTIQEKRLRVPDKFV-SKFDEB-SVAVALTVDPDGHVVRVGLRK 54

DB 6 IPSPTNKAF---IIDLSGQKSNPIPTFEIWNHFNKIQSTNMKLT-----SD 51

QY 55 ADNKIW-----PDQWQSFVDYRSIRIGVLLIFRYEGNSAFSVIFNLGSHSEINYH 106

DB 52 ASDRNWDVLDGARFAGWKDFSVSHSVRDDLLSFRHDGGMVHVSPGSRFSQILTS 111

QY 107 TGLMDSAHNHFKRARLFEDLEDAEVIPPSSVYPSLPPESTVPANKGYASSAIQTFTG 166

DB 112 SSTSDDDD---ERTVDDDEDDVDDDDNSI-----SEDDFCCKKISS----- 153

QY 167 PVKAEPTTPKIPKGRKKGNADPEE-----INSSAPRDDDDPENRSKFYESASAKR 220

DB 154 -----KKRKARETESSSDKSYLVHAHTVTPSSLLRDNCVLSKPARSNGLDR 199

QY 221 TVTAERERAINAAKTFPTNPFPRVLRVPSVLYRGCIWYLPSPGFAEKYLSIGSFIKVQ 280

DB 200 BCEIDLDRD----- 208

QY 281 LAEKOMP--VRCLYKAGRAKFSQGYEFTLENNLGEQDVCVFELLRT 325

DB 209 -HEKSWTLLLRHNKTKTQAFMRGGRWSFCRNNGIKAGSICRPFKLVS 254

Mon Jan 3 13:13:31 2005

RESULT 14
G86393
protein T24P13.6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: G86393
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Author: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86393
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-984 <STO>
A:Cross-references: UNIPROT:Q9LQV4; GB:AE005172; NID:g9295721; PIDN:AAF87027.1; GSPDB:GN
C:Genetics:
A:Gene: T24P13.6
A:Map position: 1

Query Match 7.4%; Score 133.5; DB 2; Length 984;
Best Local Similarity 21.4%; Pred. No. 0.014;
Matches 81; Conservative 46; Mismatches 151; Indels 101; Gaps 18;
QY 5 FFHKLIFSSTIOEKLRLVDPKFKVSKFDELVAVALTPDGHVWRVGLRKADNKI- 60
DB 15 FFQPII---TESRTHLNIIPVAFSKHVEGRNQNTVILR-----SDADKXTWL 60
QY 61 -----FQDGMQFVDYRISIRIGYLLIFRYEGNSAFSYYIFNLSHSYNHSTGLMDSA 113
DB 61 VKMDGLKLDGWEDFAFADHLATGDIVVFLSGEVFHTALGPSCCEIQYHT-----SS 115
QY 114 HNHFKARLPEDLEDAEVIFFSSVYPSPLPESTVPAN--KGYASSAIQTILFTGPVKA- 170
DB 116 HN-----INDDRN-----DQINIASRNSRKYKNPRKVVESLSDHSRFAKVSAM 161
QY 171 ---EEPTPTP-KIPKRGKKKNADPEEINSSAPRD-----DDPENRSKFYESA----- 215
DB 162 CLSNDRLYIPLSFARLNGLNKINSKKIYLQNEGRSWSKLVLRHDXSGMGTFFVQSGWRRFC 221
QY 216 -----SA-----RKTVTABERERAINAAKTPEPTNPFRRVVLPSYLY 254
DB 222 SENGIRQGVTFKLVKRSAPPVIRLCRAKAPKQRSV---AEYSDHSCFEGSVTPSSL- 277
QY 255 RGCIMVLPGFA-----EKVLSGISGFIKVLQAEKO---WPVRCLYKAGRAKFSQGWYEF 306
DB 278 RNDLLYLPFRSFFVNSNRLDKRCS-----EIVLKNQEGQVWVPLVLRKFKSVTYLPKGWTSF 331
QY 307 TLENLNGEGDVCVFELLRT 325
DB 332 CQVNRKAGDSFKFKLVGT 350

RESULT 15
C84639
hypothetical protein At2g24650 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: C84639
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84639
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-1440 <STO>
A:Cross-references: UNIPROT:Q9SJA0; GB:AE002093; NID:g4572680; PIDN:AAD23895.1; GSPDB:GN
C:Genetics:
A:Gene: At2g24650
A:Map position: 2

Query Match 7.4%; Score 133.5; DB 2; Length 1440;
Best Local Similarity 22.4%; Pred. No. 0.024;
Matches 87; Conservative 52; Mismatches 127; Indels 123; Gaps 20;
QY 5 FFHKLIFSSTIOEKLRLVDPKFKVSKFDELVAVALTPDGHVWRVGLRKADNKIWFQDG 64
DB 10 FFHTLVPSPF---HTHLMIPEDPFSEYIEGRSVAELKLDKDFSDKSEVKL--SDRRI--TDG 62
QY 65 WQFVDYRISIRIGYLLIFRYEGNSAFSYYIFNLSHSYNHSTGLMDSAHN-----HFKRA 120
DB 63 WEEFVWANDFRIGDVVAFRYVGN-----LVFHVSNLGNPNY-----EIEHNDLPKEKKA 112
QY 121 RLPEDLEDAEVIFFSSVYPSPLPESTVPANKGYASSAIQTILFTGPVKA-EEPTPTPKI 179
DB 113 K-----TNSEADAVSS-----SSADKSCFMAIITALDITDITLVL 149
QY 180 P-----KRGKKKNAD-----PEEINSSAPRDDDPENRSKFYESASAR----- 218
DB 150 PLHFTSANGLTRKNREIITDGGERSVLDRFD--ESSGTFYISRGWRNFCDENGQKAG 207
QY 219 -----KRTVTAE-----ERERAI----- 231
DB 208 GFFFLKLVKGGETIVLSFCPTESINGEENTREDSCSLSLDMNIVEKKYIPKPRG 267
QY 232 NAAKTPEPTN-PFPRVVLPSYLYRGCIMYLPSPGFAEKYLSGISGFIKVLQAEK---OWP 287
DB 268 SPYSSYSPSHKQFVTFTLPDYARIGKLS-LSAPFVRE--NGINKPGEICLLDKHGRKWL 324
QY 288 VRCLYKA-GRAKFSQGWYEFLENLGE 315
DB 325 TSLLLDSKGTMSLKGWKEFKVANSLETG 353

Search completed: December 30, 2004, 06:11:10
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:08:07 ; Search time 79 Seconds
(without alignments)
1548.439 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796

Sequence: 1 MPRPFHKLIFSSTIQKRL.....LLRTDFVLKVTAFRNEYV 341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1796	100.0	341	4 AAB35491	Aab35491 A thalian
2	1796	100.0	341	8 ADN73923	Adn73923 Thale cre
3	1789	99.6	341	3 AAG22414	Aag22414 Arabidops
4	1711.5	95.3	448	3 AAG54185	Aag54185 Arabidops
5	1645	91.6	377	3 AAG54186	Aag54186 Arabidops
6	1217	67.8	232	3 AAG22415	Aag22415 Arabidops
7	1217	67.8	232	3 AAG54196	Aag54196 Arabidops
8	1170	65.1	284	3 AAG54187	Aag54187 Arabidops
9	953	53.1	182	3 AAG17883	Aag17883 Arabidops
10	946	52.7	224	3 AAG58317	Aag58317 Arabidops
11	946	52.7	224	3 AAG61371	Aag61371 Arabidops
12	946	52.7	226	3 AAG61370	Aag61370 Arabidops
13	946	52.7	226	3 AAG58316	Aag58316 Arabidops
14	940	52.3	224	3 AAG17136	Aag17136 Arabidops
15	940	52.3	226	3 AAG17135	Aag17135 Arabidops
16	293.5	16.3	346	7 ADC03416	Adc03416 Rice flow
17	226.5	12.6	200	3 AAG10757	Aag10757 Arabidops
18	226.5	12.6	206	3 AAG10756	Aag10756 Arabidops
19	226.5	12.6	214	3 AAG10755	Aag10755 Arabidops
20	221	12.3	243	3 AAG54198	Aag54198 Arabidops
21	221	12.3	246	3 AAG54197	Aag54197 Arabidops
22	215.5	12.0	209	3 AAG54179	Aag54179 Arabidops
23	199	11.1	291	3 AAG47303	Aag47303 Arabidops
24	199	11.1	300	3 AAG47302	Aag47302 Arabidops
25	196	10.9	205	3 AAG54199	Aag54199 Arabidops

ALIGNMENTS

RESULT 1

AAB35491

ID AAB35491 standard; protein; 341 AA.

XX AC AAB35491;

XX AC

DT 05-NOV-2001 (first entry)

XX DE A thaliana VRN1.

XX KW VRN1; vernalisation; flowering; crop.

XX OS Arabidopsis thaliana.

XX FH Key Location/Qualifiers

FT Domain 2..94

FT Region /label= B3_DNA_binding_domain

FT Region 95..238

FT Domain /label= region_2

FT Domain 239..332

FT Domain /label= B3_DNA_binding_domain

XX WO200121822-A1.

XX 29-MAR-2001.

XX 13-SEP-2000; 2000WO-GB003525.

XX 17-SEP-1999; 99GB-00022071.

XX (PLAN-) PLANT BIOSCIENCE LTD.

XX Dean C, Levy YY;

XX WPI; 2001-273467/28.

XX N-PSDB; AAF62446.

XX Novel VRN1 polynucleotide sequence encoding a polypeptide which alters

XX vernalization response of plant in which VRN1 nucleic acid is expressed,

XX useful for influencing and assessing vernalization phenotype of plants.

XX Claim 25; Fig 7; 91pp; English.

XX The present invention provides the protein and coding sequences of

XX Arabidopsis thaliana VRN1. This protein is capable of altering the

XX vernalisation responses of a plant. Also provided are a number of PCR

XX primers used to isolate the sequences. The sequences are useful in the

XX production of crop plants, where they are able to control the timing of

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC flowering, the duration of vernalisation required, the optimum
CC temperature, or even eliminate the need for vernalisation completely. The
CC present sequence is the VRN1 protein

XX SQ Sequence 341 AA;

Query Match 100.0%; Score 1796; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.4e-167;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPRPFHKLIFSSTIOEKRLRVPDKFVSKFDLSVAVALTPDGHVVRVGLRKADNKIW 60
Db 1 MPRPFHKLIFSSTIOEKRLRVPDKFVSKFDLSVAVALTPDGHVVRVGLRKADNKIW 60
QY 61 FQDQWQEFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSHSEINHYSTGLMDSAHNHFKRA 120
Db 61 FQDQWQEFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSHSEINHYSTGLMDSAHNHFKRA 120
QY 121 RLFEDLEDEDAEVIFFPSVYSPLESTVPANKGYASSAIOTLFTGPVKABEPTTPKIP 180
Db 121 RLFEDLEDEDAEVIFFPSVYSPLESTVPANKGYASSAIOTLFTGPVKABEPTTPKIP 180
QY 181 KKRGRKKKNADPEEINSSAPRDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240
Db 181 KKRGRKKKNADPEEINSSAPRDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240
QY 241 NPFFRVLRPSLYRGCIIMYLPSGFAEKYLSGISGFIKVQLAEKQMPVRCCLYKAGRAKFS 300
Db 241 NPFFRVLRPSLYRGCIIMYLPSGFAEKYLSGISGFIKVQLAEKQMPVRCCLYKAGRAKFS 300
QY 301 QGWYEFTLENNLGECDVCVFELLRTDFVLKVTAFRVNEYV 341
Db 301 QGWYEFTLENNLGECDVCVFELLRTDFVLKVTAFRVNEYV 341

RESULT 2

ADN73923
ID ADN73923 standard; protein; 341 AA.
XX AC ADN73923;
XX DT 15-JUL-2004 (first entry)
XX DE Thale cress protein repressed in E2Fa/Dpa expressing plants Seqid 1818.
XX KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
XX KW animal feed product; thale cress; cell wall biosynthesis;
XX KW nitrogen metabolism; carbon metabolism.
XX OS Arabidopsis thaliana.
XX PN WO2004035798-A2.
XX PD 29-APR-2004.
XX PF 20-OCT-2003; 2003WO-EP011658.
XX PR 18-OCT-2002; 2002EP-00079408.
XX PA (CROP-) CROPDESIGN NV.
XX PI Inze D, De Veylder L, Vlieghe K;
XX DR WPI: 2004-348466/32.
XX DR N-PSDB; ADN73922.
XX Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
PT more proteins.
XX Claim 1; SEQ ID NO 1818; 134pp; English.

CC This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreplication, biochemistry, signal
CC transduction, storage lipid mobilisation and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polypeptide sequence is thale cress protein
CC expressed by a gene repressed 1.3 fold or more in plants overexpressing
CC the E2Fa/Dpa transcription factor, given in an exemplification of the
CC invention.

XX SQ Sequence 341 AA;

Query Match 100.0%; Score 1796; DB 8; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.4e-167;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPRPFHKLIFSSTIOEKRLRVPDKFVSKFDLSVAVALTPDGHVVRVGLRKADNKIW 60
Db 1 MPRPFHKLIFSSTIOEKRLRVPDKFVSKFDLSVAVALTPDGHVVRVGLRKADNKIW 60
QY 61 FQDQWQEFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSHSEINHYSTGLMDSAHNHFKRA 120
Db 61 FQDQWQEFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSHSEINHYSTGLMDSAHNHFKRA 120
QY 121 RLFEDLEDEDAEVIFFPSVYSPLESTVPANKGYASSAIOTLFTGPVKABEPTTPKIP 180
Db 121 RLFEDLEDEDAEVIFFPSVYSPLESTVPANKGYASSAIOTLFTGPVKABEPTTPKIP 180
QY 181 KKRGRKKKNADPEEINSSAPRDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240
Db 181 KKRGRKKKNADPEEINSSAPRDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240
QY 241 NPFFRVLRPSLYRGCIIMYLPSGFAEKYLSGISGFIKVQLAEKQMPVRCCLYKAGRAKFS 300
Db 241 NPFFRVLRPSLYRGCIIMYLPSGFAEKYLSGISGFIKVQLAEKQMPVRCCLYKAGRAKFS 300
QY 301 QGWYEFTLENNLGECDVCVFELLRTDFVLKVTAFRVNEYV 341
Db 301 QGWYEFTLENNLGECDVCVFELLRTDFVLKVTAFRVNEYV 341

RESULT 3

AAG22414
ID AAG22414 standard; protein; 341 AA.
XX AC AAG22414;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 25333.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PS 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0132866P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 27-AUG-1999; 99US-0151303P.
PR 30-AUG-1999; 99US-0151308P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.

Arabidopsis thaliana protein fragment SEQ ID NO: 69062.

Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

Arabidopsis thaliana.

RESULT 4	
AAG54185	
ID	AAG54185 standard; protein; 448 AA.
XX	
XX	
AC	AAG54185;
XX	
DT	18-OCT-2000 (first entry)
XX	

```
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140891P.
PR 30-JUN-1999; 99US-0141184P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.

PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match          95.3%; Score 1711.5; DB 3; Length 448;
Best Local Similarity 83.8%; Pred. No. 4e-159;
Matches 332; Conservative 2; Mismatches 7; Indels 55; Gaps 1;

QY 1 MRPFFHKLIFSSTIQEKKL----- 20
Db 1 MRPFFHKLIFSSTIQEKKL----- 60
QY 21 -----RVPDKFVSKFDELSVAVALTPDGHVVRGLRKADNKIWFQDGM 65
Db 61 IICFFISHTGMLVORVPDKFVSKFDELSVAVALTPDGHVVRGLRKADNKIWFQDGM 120
QY 66 QBFVDVRSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINVHSTGLMDSAHNHFKRARLPED 125
Db 121 QEFVDVRSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINVHSTGLMDSAHNHFKRARLPED 180
QY 126 LEDEDAEVIFFSSVYPSLPSTVPANKGYASSAIQTTLTGTPVKAEBPTTPKIPKQGR 185
Db 181 LEDEDAEVIFFSSVYPSLPSTVPANKGYASSAIQTTLTGTPVKAEBPTTPKIPKQGR 240
QY 186 KKKNADPEINSSAPRDDDDPENRSKFYESASARKRTVTAEERERAINAAKTPEPTNPFR 245
Db 241 KKKNADPEINSSAPRDDDDPENRSKFYESASARKRTVTAEERERAINAAKTPEPTNPFR 300
```

us-10-088-187a-11.rag

Mon Jan 3 13:13:29 2005

```
QY 246 VVLRPSYLRCIMYLPSPGFAEKYLSGSGFIKVQLAEKQMPVRCLYKAGRAKFSQGWYE 305
Db 301 VVLRPSYLRCIMYLPSPGFAEKYLSGSGFIKVQLAEKQMPVRCLYKAGRAKFSQGWYE 360
QY 306 FTLENNLGGDVCVFELLARTDFVLKVTAFRVNEV 341
Db 361 FTLENNLGGDVCVFELLARTDFVLKTYRGEANDVI 396

RESULT 5
AAG54186
ID AAG54186 standard; protein; 377 AA.
XX AC AAG54186;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 69063.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 15-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140981P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
```



```
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149310P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 91.6%; Score 1645; DB 3; Length 377;
Best Local Similarity 97.2%; Pred. No. 1.1e-152;
Matches 312; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 21 RVPDFVSKDELSSVAVALTVPDGHVWRVGLRKADNKIWFQDGHQWQEFVDYRSIRIGYLL 80
Db 5 RVPDFVSKDELSSVAVALTVPDGHVWRVGLRKADNKIWFQDGHQWQEFVDYRSIRIGYLL 64
```

```
QY 81 IFRYEGNSAFSVYIFNLHSEINHYHSTGLMDSAHNHFKEARLFEDLEDAEVIPSSVY 140
Db 65 IFRYEGNSAFSVYIFNLHSEINHYHSTGLMDSAHNHFKEARLFEDLEDAEVIPSSVY 124
QY 141 PSPLPESTVPANKGYASSAIQTLFTGPVKAEBPTTPKIPKGRGRKKQKQADPEEINSSAP 200
Db 125 PSPLPESTVPANKGYASSAIQTLFTGPVKAEBPTTPKIPKGRGRKKQKQADPEEINSSAP 184
QY 201 RDDDPENRSKFYESASARKRTVTAEERERAINAAKTFFETNPFFRVVLRPSLYRCIMY 260
Db 185 RDDDPENRSKFYESASARKRTVTAEERERAINAAKTFFETNPFFRVVLRPSLYRCIMY 244
QY 261 LPSGFAEKYLSGISGFIKVQLAEKQMPVRCLYKAGRAKFSQGWYFTLENNLGEQDVCVF 320
Db 245 LPSGFAEKYLSGISGFIKVQLAEKQMPVRCLYKAGRAKFSQGWYFTLENNLGEQDVCVF 304
QY 321 ELLRTRDFVLKVTAFRVNEYV 341
Db 305 ELLRTRDFVLKTYRGEAVDYI 325
```

RESULT 6

AAG22415

ID AAG22415 standard; protein; 232 AA.

XX

AC AAG22415;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 25334.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132407P.

PR 05-MAY-1999; 99US-0132484P.

PR 06-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 18-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143824P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.

PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 67.8%; Score 1217; DB 3; Length 232;
Best Local Similarity 100.0%; Pred. No. 6.4e-111;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 MDSAHNHFKRARLFEDEDEDAVIPPSSVYPSLPSTVPANKGVASSAIOITLFTGPVK 169
DB 1 MDSAHNHFKRARLFEDEDEDAVIPPSSVYPSLPSTVPANKGVASSAIOITLFTGPVK 60

QY 170 ABEPTPTPKPKKGRKKKNADPEEINSSAPRDDDDPENRSKFYESASAKRTVTAEER 229
DB 61 ABEPTPTPKPKKGRKKKNADPEEINSSAPRDDDDPENRSKFYESASAKRTVTAEER 120

QY 230 AINAATFPTNPFFRVLRPSLYRGCIIMYLPSPGAEKYLSGISGFIKVQLAEKQMPVR 289
DB 121 AINAATFPTNPFFRVLRPSLYRGCIIMYLPSPGAEKYLSGISGFIKVQLAEKQMPVR 180

QY 290 CLYKAGRAKPSQGWYEFTLENNLGECDVCVFELLRTDFVLKVTAFRVNEYV 341
DB 181 CLYKAGRAKPSQGWYEFTLENNLGECDVCVFELLRTDFVLKVTAFRVNEYV 232

RESULT 7
AAG54196
ID AAG54196 standard; protein; 232 AA.
AC AAG54196;
XX
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 69077.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2008EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132485P.
PR 05-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.

14-MAY-1999; 99US-0134219P.
14-MAY-1999; 99US-0134221P.
14-MAY-1999; 99US-0134370P.
18-MAY-1999; 99US-0134768P.
19-MAY-1999; 99US-0134941P.
20-MAY-1999; 99US-0135124P.
21-MAY-1999; 99US-0135353P.
24-MAY-1999; 99US-0135629P.
25-MAY-1999; 99US-0136021P.
27-MAY-1999; 99US-0136392P.
28-MAY-1999; 99US-0136782P.
01-JUN-1999; 99US-0137222P.
03-JUN-1999; 99US-0137528P.
04-JUN-1999; 99US-0137502P.
07-JUN-1999; 99US-0137724P.
08-JUN-1999; 99US-0138094P.
10-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
14-JUN-1999; 99US-0139119P.
16-JUN-1999; 99US-0139452P.
16-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139457P.
18-JUN-1999; 99US-0139458P.
18-JUN-1999; 99US-0139459P.
18-JUN-1999; 99US-0139460P.
18-JUN-1999; 99US-0139461P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
18-JUN-1999; 99US-0139763P.
21-JUN-1999; 99US-0139817P.
22-JUN-1999; 99US-0139899P.
23-JUN-1999; 99US-0140353P.
23-JUN-1999; 99US-0140354P.
24-JUN-1999; 99US-0140695P.
28-JUN-1999; 99US-0140821P.
29-JUN-1999; 99US-0140991P.
30-JUN-1999; 99US-0141287P.
01-JUL-1999; 99US-0141842P.
01-JUL-1999; 99US-0142154P.
02-JUL-1999; 99US-0142055P.
06-JUL-1999; 99US-0142390P.
08-JUL-1999; 99US-0142803P.
09-JUL-1999; 99US-0142920P.
12-JUL-1999; 99US-0142977P.
13-JUL-1999; 99US-0143542P.
14-JUL-1999; 99US-0143624P.
15-JUL-1999; 99US-0144005P.
16-JUL-1999; 99US-0144085P.
16-JUL-1999; 99US-0144086P.
19-JUL-1999; 99US-0144325P.
19-JUL-1999; 99US-0144331P.
19-JUL-1999; 99US-0144332P.
19-JUL-1999; 99US-0144333P.
19-JUL-1999; 99US-0144334P.
19-JUL-1999; 99US-0144335P.
20-JUL-1999; 99US-0144352P.
20-JUL-1999; 99US-0144632P.
20-JUL-1999; 99US-0144884P.
21-JUL-1999; 99US-0144814P.
21-JUL-1999; 99US-0145086P.
21-JUL-1999; 99US-0145088P.
22-JUL-1999; 99US-0145085P.
22-JUL-1999; 99US-0145087P.
22-JUL-1999; 99US-0145089P.
22-JUL-1999; 99US-0145192P.
23-JUL-1999; 99US-0145145P.
23-JUL-1999; 99US-0145218P.
23-JUL-1999; 99US-0145224P.

26-JUL-1999;	99US-01452769;
27-JUL-1999;	99US-0145913P;
27-JUL-1999;	99US-0145918P;
27-JUL-1999;	99US-0145919P;
28-JUL-1999;	99US-0145951P;
02-AUG-1999;	99US-0146386P;
02-AUG-1999;	99US-0146388P;
02-AUG-1999;	99US-0146389P;
03-AUG-1999;	99US-0147038P;
04-AUG-1999;	99US-0147204P;
04-AUG-1999;	99US-0147302P;
05-AUG-1999;	99US-0147192P;
05-AUG-1999;	99US-0147260P;
06-AUG-1999;	99US-0147303P;
06-AUG-1999;	99US-0147416P;
09-AUG-1999;	99US-0147493P;
09-AUG-1999;	99US-0147935P;
10-AUG-1999;	99US-0148171P;
11-AUG-1999;	99US-0148319P;
12-AUG-1999;	99US-0148341P;
13-AUG-1999;	99US-0148465P;
13-AUG-1999;	99US-0148564P;
16-AUG-1999;	99US-0149368P;
17-AUG-1999;	99US-0149175P;
18-AUG-1999;	99US-0149426P;
20-AUG-1999;	99US-0149722P;
20-AUG-1999;	99US-0149723P;
20-AUG-1999;	99US-0149929P;
23-AUG-1999;	99US-0149902P;
23-AUG-1999;	99US-0149930P;
25-AUG-1999;	99US-0150566P;
25-AUG-1999;	99US-0150884P;
26-AUG-1999;	99US-0151065P;
27-AUG-1999;	99US-0151066P;
27-AUG-1999;	99US-0151080P;
27-AUG-1999;	99US-0151303P;
30-AUG-1999;	99US-0151438P;
31-AUG-1999;	99US-0151930P;
01-SEP-1999;	99US-0152363P;
07-SEP-1999;	99US-0153070P;
10-SEP-1999;	99US-0153758P;
13-SEP-1999;	99US-0154018P;
16-SEP-1999;	99US-0154039P;
16-SEP-1999;	99US-0154779P;
20-SEP-1999;	99US-0155133P;
22-SEP-1999;	99US-0155486P;
23-SEP-1999;	99US-0155659P;
24-SEP-1999;	99US-0156458P;
28-SEP-1999;	99US-0156596P;
29-SEP-1999;	99US-0157117P;
04-OCT-1999;	99US-0157753P;
05-OCT-1999;	99US-0157865P;
06-OCT-1999;	99US-0158292P;
07-OCT-1999;	99US-0158232P;
08-OCT-1999;	99US-0158369P;
12-OCT-1999;	99US-0159293P;
13-OCT-1999;	99US-0159294P;
13-OCT-1999;	99US-0159329P;
14-OCT-1999;	99US-0159330P;
14-OCT-1999;	99US-0159331P;
14-OCT-1999;	99US-0159332P;
21-OCT-1999;	99US-0160776P;
21-OCT-1999;	99US-0160768P;
21-OCT-1999;	99US-0160770P;
21-OCT-1999;	99US-0160814P;
21-OCT-1999;	99US-0160853P;
18-OCT-1999;	99US-0159584P;
21-OCT-1999;	99US-0160741P;
21-OCT-1999;	99US-0160767P;
21-OCT-1999;	99US-0160768P;
22-OCT-1999;	99US-0160770P;
22-OCT-1999;	99US-0160981P;
22-OCT-1999;	99US-0160989P;
22-OCT-1999;	99US-0160989P;

PR 06-MAY-1999; 99US-0132486P. 22-JUL-1999; 99US-0145089P.
PR 06-MAY-1999; 99US-0132487P. 22-JUL-1999; 99US-0145192P.
PR 07-MAY-1999; 99US-0132863P. 23-JUL-1999; 99US-0145145P.
PR 11-MAY-1999; 99US-0134256P. 23-JUL-1999; 99US-0145218P.
PR 14-MAY-1999; 99US-0134218P. 23-JUL-1999; 99US-0145224P.
PR 14-MAY-1999; 99US-0134219P. 26-JUL-1999; 99US-0145276P.
PR 14-MAY-1999; 99US-0134221P. 27-JUL-1999; 99US-0145913P.
PR 14-MAY-1999; 99US-0134370P. 27-JUL-1999; 99US-0145918P.
PR 18-MAY-1999; 99US-01344768P. 27-JUL-1999; 99US-0145951P.
PR 19-MAY-1999; 99US-0134941P. 28-JUL-1999; 99US-0145951P.
PR 20-MAY-1999; 99US-0135124P. 02-AUG-1999; 99US-0146386P.
PR 21-MAY-1999; 99US-0135353P. 02-AUG-1999; 99US-0146388P.
PR 24-MAY-1999; 99US-0135629P. 02-AUG-1999; 99US-0146389P.
PR 25-MAY-1999; 99US-0136021P. 03-AUG-1999; 99US-0147038P.
PR 27-MAY-1999; 99US-0136392P. 04-AUG-1999; 99US-0147204P.
PR 28-MAY-1999; 99US-0136782P. 04-AUG-1999; 99US-0147302P.
PR 01-JUN-1999; 99US-0137222P. 05-AUG-1999; 99US-0147192P.
PR 03-JUN-1999; 99US-0137528P. 05-AUG-1999; 99US-0147260P.
PR 04-JUN-1999; 99US-0137502P. 06-AUG-1999; 99US-0147303P.
PR 07-JUN-1999; 99US-0137724P. 06-AUG-1999; 99US-0147416P.
PR 08-JUN-1999; 99US-0138094P. 09-AUG-1999; 99US-0147493P.
PR 10-JUN-1999; 99US-0138540P. 09-AUG-1999; 99US-0147935P.
PR 10-JUN-1999; 99US-0138847P. 10-AUG-1999; 99US-0148171P.
PR 14-JUN-1999; 99US-0139119P. 11-AUG-1999; 99US-0148315P.
PR 16-JUN-1999; 99US-0139452P. 12-AUG-1999; 99US-0148341P.
PR 16-JUN-1999; 99US-0139453P. 13-AUG-1999; 99US-0148565P.
PR 17-JUN-1999; 99US-0139492P. 13-AUG-1999; 99US-0148684P.
PR 18-JUN-1999; 99US-0139454P. 16-AUG-1999; 99US-0149368P.
PR 18-JUN-1999; 99US-0139455P. 17-AUG-1999; 99US-0149175P.
PR 18-JUN-1999; 99US-0139456P. 18-AUG-1999; 99US-0149428P.
PR 18-JUN-1999; 99US-0139457P. 20-AUG-1999; 99US-0149722P.
PR 18-JUN-1999; 99US-0139458P. 20-AUG-1999; 99US-0149723P.
PR 18-JUN-1999; 99US-0139459P. 23-AUG-1999; 99US-0149929P.
PR 18-JUN-1999; 99US-0139460P. 23-AUG-1999; 99US-0149902P.
PR 18-JUN-1999; 99US-0139461P. 25-AUG-1999; 99US-0149930P.
PR 18-JUN-1999; 99US-0139462P. 26-AUG-1999; 99US-0150868P.
PR 18-JUN-1999; 99US-0139463P. 26-AUG-1999; 99US-0150884P.
PR 18-JUN-1999; 99US-0139750P. 27-AUG-1999; 99US-0151065P.
PR 18-JUN-1999; 99US-0139763P. 27-AUG-1999; 99US-0151066P.
PR 21-JUN-1999; 99US-0139817P. 27-AUG-1999; 99US-0151080P.
PR 22-JUN-1999; 99US-0139899P. 30-AUG-1999; 99US-0151303P.
PR 23-JUN-1999; 99US-0140353P. 31-AUG-1999; 99US-0151438P.
PR 23-JUN-1999; 99US-0140354P. 01-SEP-1999; 99US-0151930P.
PR 24-JUN-1999; 99US-0140695P. 07-SEP-1999; 99US-0152363P.
PR 28-JUN-1999; 99US-0140823P. 10-SEP-1999; 99US-0153070P.
PR 29-JUN-1999; 99US-0140991P. 13-SEP-1999; 99US-0153758P.
PR 30-JUN-1999; 99US-0141287P. 15-SEP-1999; 99US-0154018P.
PR 01-JUL-1999; 99US-0141842P. 16-SEP-1999; 99US-0154039P.
PR 01-JUL-1999; 99US-0142154P. 20-SEP-1999; 99US-0154779P.
PR 02-JUL-1999; 99US-0142055P. 22-SEP-1999; 99US-0155139P.
PR 06-JUL-1999; 99US-0142390P. 23-SEP-1999; 99US-0155486P.
PR 08-JUL-1999; 99US-0142803P. 24-SEP-1999; 99US-0155659P.
PR 09-JUL-1999; 99US-0142920P. 28-SEP-1999; 99US-0156458P.
PR 12-JUL-1999; 99US-0142977P. 29-SEP-1999; 99US-0156596P.
PR 13-JUL-1999; 99US-0143542P. 04-OCT-1999; 99US-0157117P.
PR 14-JUL-1999; 99US-0143624P. 05-OCT-1999; 99US-0157753P.
PR 15-JUL-1999; 99US-0144005P. 06-OCT-1999; 99US-0157865P.
PR 16-JUL-1999; 99US-0144085P. 07-OCT-1999; 99US-0158029P.
PR 16-JUL-1999; 99US-0144086P. 08-OCT-1999; 99US-0158232P.
PR 19-JUL-1999; 99US-0144325P. 12-OCT-1999; 99US-0158369P.
PR 19-JUL-1999; 99US-0144331P. 13-OCT-1999; 99US-0159293P.
PR 19-JUL-1999; 99US-0144332P. 13-OCT-1999; 99US-0159294P.
PR 19-JUL-1999; 99US-0144333P. 13-OCT-1999; 99US-0159295P.
PR 19-JUL-1999; 99US-0144334P. 14-OCT-1999; 99US-0159329P.
PR 19-JUL-1999; 99US-0144335P. 14-OCT-1999; 99US-0159330P.
PR 20-JUL-1999; 99US-0144352P. 14-OCT-1999; 99US-0159331P.
PR 20-JUL-1999; 99US-0144632P. 14-OCT-1999; 99US-0159637P.
PR 20-JUL-1999; 99US-0144684P. 18-OCT-1999; 99US-0159638P.
PR 21-JUL-1999; 99US-0144814P. 21-OCT-1999; 99US-0159584P.
PR 21-JUL-1999; 99US-0145086P. 21-OCT-1999; 99US-0160741P.
PR 21-JUL-1999; 99US-0145088P. 21-OCT-1999; 99US-0160767P.
PR 22-JUL-1999; 99US-0145085P. 21-OCT-1999; 99US-0160768P.
PR 22-JUL-1999; 99US-0145087P. 21-OCT-1999; 99US-0160770P.

PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161408P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161408P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 65.1%; Score 1170; DB 3; Length 284;
Best Local Similarity 96.1%; Pred. No. 3.6e-106;
Matches 223; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 110 MDSAHNHFRAFLFDEDEDAEVFPSSVYPSLPSTVPANKGYASSAIQTLTGPVK 169
Db 1 MDSAHNHFRAFLFDEDEDAEVFPSSVYPSLPSTVPANKGYASSAIQTLTGPVK 60

QY 170 ABEPTPTPKIPKRGKKNADPEEINGSAPRDDDPENRSKFYESASARKRTVTAAERER 229
Db 61 ABEPTPTPKIPKRGKKNADPEEINGSAPRDDDPENRSKFYESASARKRTVTAAERER 120

QY 230 AINAAKTFEPTNPFPRVLRPSLYRGCMYLPSCGFAEYLSGSGFKVQLAEKQWVR 289
Db 121 AINAAKTFEPTNPFPRVLRPSLYRGCMYLPSCGFAEYLSGSGFKVQLAEKQWVR 180

QY 290 CLYKAGRAKFSQGWTFEFTLENNLGECDVCVFELLTRDVLKVTAFRNEVY 341
Db 181 CLYKAGRAKFSQGWTFEFTLENNLGECDVCVFELLTRDVLKVTAFRNEVY 232

RESULT 9
AAG17883
XX ID AAG17883 standard; protein; 182 AA.
XX AC
XX AAG17883;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 19076.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142053P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.

```
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 28-JUL-1999; 99US-0145319P.
PR 28-JUL-1999; 99US-0145351P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.

Query Match 53.1%; Score 953; DB 3; Length 182;
Best Local Similarity 100.0%; Pred. No. 4e-85; 0; Indels 0; Gaps 0;
Matches 181; Conservative 0; Mismatches 0;

QY 161 QTLFTGPVKAERTPTPKIPKRGKKKNADPEEINSSAPRDDDDPENRSKPFYESASARKR 220
Db |||||
QY 2 QTLFTGPVKAERTPTPKIPKRGKKKNADPEEINSSAPRDDDDPENRSKPFYESASARKR 61
Db |||||
QY 221 TVTAEERERAINAAKTFTPTNPFVRLRPSLYRGCMYLPSPGFAEKYLSGISGFIKVQ 280
Db |||||
QY 281 LAEKQMPVRCLYKAGRAKESQGWYFTLENNLGEQDVCVFELLRTDFVLKVTAFRVNEY 340
Db |||||
QY 122 LAEKQMPVRCLYKAGRAKESQGWYFTLENNLGEQDVCVFELLRTDFVLKVTAFRVNEY 181
QY 341 V 341
Db 182 V 182

RESULT 10
AAG58317
ID AAG58317 standard; protein; 224 AA.
XX
AC AAG58317;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 75264.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
```



```
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160860P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 52.7%; Score 946; DB 3; Length 224;
Best Local Similarity 81.2%; Pred. No. 2.6e-94;
Matches 190; Conservative 14; Mismatches 18; Indels 12; Gaps 5;

QY 110 MSAHNP-KRRLFDLDEDAEVIFFPSVVPSPLESTVPANKY-ASSAIQTLFTGP 167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MDSAQQFNKRRLFDPELDAKVIYPSN-----PESTEPVNGYGGSTAISPFK-E 53

QY 168 VKAEETPTPKTKGGRKKQADPEINSAPRDDDDPENRSKFYESASARKRTVTAER 227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54 SKABE---TPKVLKKGKQKQNPNEVNSSTFGDDSDENRSKFYESASARKRTVTAER 110

QY 228 ERAINAAKTFEPTNPFRVRLPSYLYRGICIMYLPSPGFAKYLSGISGFIKVQLAEKOMP 287
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111 ERAVNAAKTFEPTNPFRVRLPSYLYRGICIMYLPSPGFAKYLSGISGFIKVQLAEKOMP 170

QY 288 VRCLYKAGRAKFSQGYEFTLENNLGEQVCFPELLRTDFVLKVTAFRVNEYV 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 171 VRCLYKAGRAKFSQGYEFTLENNLGEQVCFPELLRTDFVLKVTAFRVNEYV 224

RESULT 11
AAG61371
ID AAG61371 standard; protein; 224 AA.
XX
AC AAG61371;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 79595.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000BP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137503P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142053P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
```


XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147034P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153707P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.

PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match .52.7%; Score 946; DB 3; Length 226;

Best Local Similarity 81.2%; Pred. No. 2,7e-84;
Matches 190; Conservative 14; Mismatches 18; Indels 12; Gaps 5;

QY 110 MDSAHNF-KRRLFEDEDEDAVIFSPVSPPLPSTVPANKY-ASSAIQTLFTGP 167
DB 3 MDSAQNFKRARLFEDPELKDVKYFSN-----PESTEPVNGYGGSTAIQSFYK-E 55
QY 168 VKAEEPTPKPKPKRGRKKVADDEEINSSAPRDDDENRSKFYESASARKRTVTABER 227
DB 56 SKAEE---TPKVLKRRKKKNPNPEEVNSSTPGDSDENRSKFYESASARKRTVTABER 112
QY 228 ERAINAAKTFEPTNPFRRVLRPSLYRGICINYLPSGFAEKYLSGISGFIKVLAEKQWP 287
DB 113 ERAVNAAKTFEPTNPFRRVLRPSLYRGICINYLPSGFAEKYLSGISGFIKVLGEKQWP 172
QY 288 VRCLYKAGRAKFSQGWYEFTLENNIGEGDVCFVLLRTRDFVLKTAFRVNEYV 341
DB 173 VRCLYKAGRAKFSQGWYEFTLENNIGEGDVCFVLLRTRDFVLEVTAFRVNEYV 226

RESULT 13

AAG58316
ID AAG58316 standard; protein; 226 AA.

XX AC AAG58316;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 75263.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0123788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140832P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.

```
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158363P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 52.7%; Score 946; DB 3; Length 226;
Best Local Similarity 81.2%; Pred. No. 2.7e-84;
Matches 190; Conservative 14; Mismatches 18; Indels 12; Gaps 5;

QY 110 MDSAHNF-KCARLEFDEDEDAVIFPSSVYPSLPSTVTPANKY-ASSAIQTLTGP 167
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3 MDSAQNQFNKARLFEDPELKDAKVIYPSN-----PESTEPVNKGYGGSTAIOGFFK-E 55
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 168 VKABEPTTPKIPKGRKKKNADPEEINSSAPRDDDPENRSKFYESASARKRTVTAER 227
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 56 SKAEE---TPKVLKGRKKKNPNPEEVNNSSTPGDDSDENRSKFYESASARKRTVTAER 112
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 228 ERAINAAKTFTPTNPFPRVLRPSYLYRGCMYLPSPGFAEKYLSIGSFIKVLAEKOMP 287
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 113 ERAVNAAKTFTPTNPFPRVLRPSYLYRGCMYLPSPGFAEKYLSIGSFIKVLAEKOMP 172
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 288 VRCLYKAGRAKFSQGWYFTLENNIGEDVCVFELLRTDRFVLKTAFRVNEYV 341
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 173 VRCLYKAGRAKFSQGWYFTLENNIGEDVCVFELLRTDRFVLKTAFRVNEYV 226
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
AAG17136
ID AAG17136 standard; protein; 224 AA.
XX
AC AAG17136;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18043.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
```

hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX Arabidopsis thaliana.
XX EP1033405-A2.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134258P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147352P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.

PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161320P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 52.3%; Score 940; DB 3; Length 224;		
Best Local Similarity 80.8%; Pred. No. 1e-83;		
Matches 189; Conservative 14; Mismatches 19; Indels 12; Gaps 5;		
Qy	110	MDSAHNR-KRALPEDLEDAEVIFFSSVYPSLPPESTVPANKGY-ASSAIQTLFTGP 167
Db	1	MDSAQNQFNKRALFEDPELKDAKIYPSN-----PEFEPNKGYGSGTAIQSFFK-E 53
Qy	168	VKAEEPTTPPKIKRGRKKKNADPEEINSSAPRDDDDPENRSKSFYESASARKRTVTABER 227
Db	54	SKAEE---TPKVLKKRKKKNPNPEEVNNSSTPGGDDSENRSKSFYESASARKRTVTABER 110
Qy	228	ERAINAAKTFFETNFFRVLPSPSYLYRCIMYLPSPGFAEKYLSGISPIKVLAKQWP 287
Db	111	ERAVNAAKTFFETNFFRVLPSPSYLYRCIMYLPSPGFAEKYLSGISPIKVLAKQWP 170
Qy	288	VRCLYKAGRAKFSQGWYEFTELENNIGEGDVCVFELLRTDFVLKVTAFRVNEYV 341
Db	171	VRCLYKAGRAKFSQGWYEFTELENNIGEGDVCVFELLRTDFVLKVTAFRVNEYV 224
RESULT 15		
AAG17135		
ID	AAG17135 standard; protein; 226 AA.	
XX		
AC	AAG17135;	
XX		

DT	17-OCT-2000	(first entry)
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 18042.	
XX	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX	Arabidopsis thaliana.	
OS	Arabidopsis thaliana.	
XX	EP1033405-A2.	
PN	06-SEP-2000.	
XX		
PF	25-FEB-2000;	2000EP-00301439.
PR	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.

Job time : 82 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 06:07:48 ; Search time 25 Seconds
(without alignments)
904.578 Million cell updates/sec

Title: US-10-088-187A-11
Perfect score: 1796
Sequence: 1 MRPFPFKLIFSSTIQEKRL.....LLRTRDFVLKVTAFRNVEYV 341

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/prodata/1/1aa/ECTUS-COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97.5	5.4	1178	4	US-09-543-681A-7819
2	97.5	5.4	8991	4	US-08-714-741-32
3	97	5.4	399	4	US-09-107-532A-6398
4	95.5	5.3	834	3	US-08-539-205A-6
5	95.5	5.3	834	4	US-09-392-163A-6
6	94.5	5.3	497	4	US-09-248-796A-15909
7	94.5	5.3	693	4	US-09-248-796A-26890
8	93.5	5.2	274	4	US-09-489-039A-10361
9	92.5	5.2	194	4	US-08-529-055-64
10	92.5	5.2	496	3	US-08-740-223A-16
11	92.5	5.2	496	4	US-09-709-188-16
12	92	5.1	748	4	US-09-270-767-4471
13	91.5	5.1	1736	4	US-09-919-497-98
14	90	5.0	1061	4	US-09-762-481B-2
15	88.5	4.9	286	3	US-08-960-507-20
16	88.5	4.9	286	4	US-09-136-801-20
17	88.5	4.9	286	4	US-09-202-088A-20
18	88.5	4.9	334	4	US-09-248-796A-21169
19	88.5	4.9	480	3	US-08-740-223A-8
20	88.5	4.9	480	4	US-09-709-188-8
21	88.5	4.9	496	1	US-08-373-579-6
22	88.5	4.9	496	2	US-08-418-595-6
23	88.5	4.9	496	2	US-08-665-926-6
24	88.5	4.9	496	3	US-09-162-437-6
25	88.5	4.9	496	3	US-08-740-223A-6
26	88.5	4.9	496	3	US-09-351-457-4
27	88.5	4.9	496	3	US-09-561-500-4

28	88.5	4.9	496	3	US-09-561-108-4	Sequence 4, Appl
29	88.5	4.9	496	4	US-09-351-543-4	Sequence 4, Appl
30	88.5	4.9	496	4	US-09-561-526-4	Sequence 4, Appl
31	88.5	4.9	496	4	US-09-202-491-8	Sequence 8, Appl
32	88.5	4.9	496	4	US-08-817-318-6	Sequence 6, Appl
33	88.5	4.9	496	4	US-09-709-188-6	Sequence 6, Appl
34	88.5	4.9	496	4	US-09-561-499-4	Sequence 4, Appl
35	88.5	4.9	496	4	US-09-442-717-6	Sequence 6, Appl
36	88.5	4.9	496	4	US-09-689-020-6	Sequence 6, Appl
37	88.5	4.9	496	4	US-09-998-831-4	Sequence 4, Appl
38	88.5	4.9	496	4	US-09-561-005-4	Sequence 4, Appl
39	88.5	4.9	496	4	US-09-819-386-4	Sequence 4, Appl
40	88.5	4.9	498	3	US-08-740-223A-20	Sequence 20, Appl
41	88.5	4.9	498	4	US-09-709-188-20	Sequence 20, Appl
42	88.5	4.9	499	3	US-08-740-223A-24	Sequence 24, Appl
43	88.5	4.9	499	4	US-09-709-188-24	Sequence 24, Appl
44	87.5	4.9	389	4	US-09-248-796A-17225	Sequence 17225, A
45	87.5	4.9	432	4	US-09-248-796A-18913	Sequence 18913, A

ALIGNMENTS

RESULT 1

US-09-543-681A-7819
; Sequence 7819, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7819
; LENGTH: 1178
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7819

Query Match		5.4%	Score 97.5;	DB 4;	Length 1178;
Best Local Similarity		21.7%	Pred. No. 0.47;		
Matches		53;	Conservative	40;	Mismatches 98;
				Indels	53;
				Gaps	10;
QY	15	IOEKRLR-----	VDPKFVSKPKDELVAVALTPDGHVVRVGLRKADNKKIWFQDGM	65	
DB	432	IEEALKENTHEVHAIVPQIASYLLNEKRAVT-----	DIELRQANVKV-----	476	
QY	66	QEFVDYRIGYLLIFRYEGNSAFVIFNLHSEINHYHSTGLMDSAHNHFARLFED	125		
DB	477	-VIVPNDQMQTFPHFVIRKGEIISLSYNLA-----	QYHETQLNEADEAVPERKAPQ	531	
QY	126	-----LEDEDARVIFPSSVYVSP----	LPESTVPANKGYASSAQTTLFTG-----	PV	168
DB	532	PAISAFALQDINSADTNNKKSAPTAKQVKSETTENGPF-PSAITKFSNLPASQPEPV	590		
QY	169	KAEPTPTPKPKKKKKKNADPEEINSSAPRDDDPNRSKFYSASARKKVTTABERE	228		
DB	591	VEEKQESKKNENRRRRNNRRNNRRDND---RNRDHE---	RNRDVTSEENR	644	
QY	229	RAIN 232			
DB	645	KVRN 648			

RESULT 2

US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:

APPLICANT: Briles, David E.
 APPLICANT: McDaniel, Larry S.
 APPLICANT: Swiatlo, Edwin
 APPLICANT: Yother, Janet
 APPLICANT: Crain, Marilyn J.
 APPLICANT: Hollingshead, Susan
 APPLICANT: Tart, Rebecca
 APPLICANT: Brooks-Walter, Alexis
 TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
 TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
 TITLE OF INVENTION: PORTIONS AND PRODUCTS
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Curtis, Morris & Safford, P.C.
 STREET: 530 Fifth Avenue
 CITY: New York
 STATE: New York
 COUNTRY: U.S.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/714,741
 FILING DATE: 16-SEP-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Frommer Esq., William S.
 REGISTRATION NUMBER: 25,506
 REFERENCE/DOCKET NUMBER: 454312-2460
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 840-3333
 TELEFAX: (212) 840-0712
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8991 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: amino acid
 US-08-714-741-32

Query Match 5.4%; Score 97.5; DB 4; Length 8991;
 Best Local Similarity 24.4%; Pred.No.12;
 Matches 38; Conservative 17; Mismatches 72; Indels 29; Gaps 4;

QY	125	DLEDDAEVIFSSVTPSLPSTVPANKGYASSAQTLTGTPVKAEEPTPTKIPKKRG	184
Db	5678	DREETPAPQPPQAPAPAKPEQAPAP-----PKPEQAPAPAK-PEQPA	5720
QY	185	RKKKNADPEEINSSAPRDDDPENRSFYESASARKRTVTAERERAINAAKTPEPTNPF	244
Db	5721	PAPKPPQ-----APAKPEQAPAKPEPAEPTQPEKPAFTKTVRLKVAE-----F	5768
QY	245	RVVLRPSYLYRCIMYLPSPGFAEKYLSGISGFIVQ	280
Db	5769	GVQLRDAGGSNNVGAYFKEGLEETTAEXEAGLGKAE	5804

RESULT 3
 US-09-107-532A-6398
 Sequence 6398, Application US/09107532A
 Patent No. 6583275
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street

CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107.532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Denise
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6398:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:

NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...399
SEQUENCE DESCRIPTION: SEQ ID NO: 6398:

US-09-107-532A-6398

Query Match 5.4%; Score 97; DB 4; Length 399;
Best Local Similarity 28.9%; Pred. No. 0.095;
Matches 37; Conservative 14; Mismatches 71; Indels 6; Gaps 5;

QY 118 KRALFDLDRDAEVIFPSSVPSPLPESTVPANKGYASSAIQTILFTGPKVAEPTTP 177
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
75 KQRTAPEMAAEQTEPLLTPLEYLADSTAEPFPDPAAMRGAVDMPLLDAP I-AEEPTPAP 133

QY 178 --KIPKKRGGKKKNADPEINSAPRDDDPENRS-KFYESASARKTIVTAERERAINAA 234
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |

DB 134 VRKPQQRKGKKKKKOSTKTFTEDARPGSEKTMRGVEGAPAKRR-TSEGSR-LQFD 191
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |

QY 235 KTFFETNP 242
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |

DB 192 KEPEPADP 199

RESULT 4
US-08-539-205A-6
Sequence 6, Application US/08539205A
Patent No. 6001619
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Caligiuri, Maureen
APPLICANT: Neifsky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-539-205A-6

```

```

Query Match          5.3%; Score 95.5; DB 3; Length 834;
Best Local Similarity 21.0%; Pred. No. 0.45;
Matches 66; Conservative 32; Mismatches 109; Indels 107; Gaps 15;

QY 1 MRPFFPKLIFSSTIQEKRLRPDPKFKVSKFDLSVAVALTVPDG-----H 46
Db 1 MRPY-----TFKDFLLR-PRSHKSRVKGFLKXAYMPKNGQDENSDQDDMEH 51

QY 47 VWRVGLRKADNKKWFQD-----GWQEFVDYRSIRIGYLLIFRYEGNSAFSVYFNL 97
Db 52 GWEV-VDSNDSASQHQEELPPPLPPGWEEKVDNLG-----RTYYVNH 93

QY 98 SHSEINYHSTGLMD-----SAHHFK-RARLFEDLEDEDAEVIFPSSVYP 141
Db 94 NNRTQWHRPSLMDVSSSDNNIRQINQEAHRRFRSRHISEDLEPE-----PSEGDD 147

QY 142 SPLPESTVPANKGYASSAIQTLFTGPKVKAEEPTPKPKKGRKKGNADPEEINSSAPR 201
Db 148 VPEPWETISEEVNIAGDSLGLAL-----PPPVSPGSR-----TSPQELSEELSR 192

QY 202 -----DDDPENRSKPYE-SASARKRTVTAERERAINAAKTFTPTNPFVVLRLPSVLY 254
Db 193 RLQITPDSNGEQFSLLIQREPSRLRSCSVTD---AAVEQGHLP-----PSVAY 239

QY 255 RGCIMVLPSPGPAEK 268
Db 240 VHTTGLPSGWEER 253

```

```

RESULT 5
US-09-392-163A-6
; Sequence 6, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-392-163A-6

```

```

Query Match          5.3%; Score 95.5; DB 4; Length 834;
Best Local Similarity 21.0%; Pred. No. 0.45;
Matches 66; Conservative 32; Mismatches 109; Indels 107; Gaps 15;

QY 1 MRPFFPKLIFSSTIQEKRLRPDPKFKVSKFDLSVAVALTVPDG-----H 46
Db 1 MRPY-----TFKDFLLR-PRSHKSRVKGFLKXAYMPKNGQDENSDQDDMEH 51

QY 47 VWRVGLRKADNKKWFQD-----GWQEFVDYRSIRIGYLLIFRYEGNSAFSVYFNL 97
Db 52 GWEV-VDSNDSASQHQEELPPPLPPGWEEKVDNLG-----RTYYVNH 93

QY 98 SHSEINYHSTGLMD-----SAHHFK-RARLFEDLEDEDAEVIFPSSVYP 141
Db 94 NNRTQWHRPSLMDVSSSDNNIRQINQEAHRRFRSRHISEDLEPE-----PSEGDD 147

QY 142 SPLPESTVPANKGYASSAIQTLFTGPKVKAEEPTPKPKKGRKKGNADPEEINSSAPR 201
Db 148 VPEPWETISEEVNIAGDSLGLAL-----PPPVSPGSR-----TSPQELSEELSR 192

QY 202 -----DDDPENRSKPYE-SASARKRTVTAERERAINAAKTFTPTNPFVVLRLPSVLY 254
Db 193 RLQITPDSNGEQFSLLIQREPSRLRSCSVTD---AAVEQGHLP-----PSVAY 239

QY 255 RGCIMVLPSPGPAEK 268
Db 240 VHTTGLPSGWEER 253

```

```

RESULT 6
US-09-248-796A-15909
; Sequence 15909, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15909
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-15909

```

[illegible]

REGISTRATION NUMBER: 25,506
REFERENCES/DOCKET NUMBER: 454312-2400
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-529-055-64

Query Match 5.2%; Score 92.5; DB 4; Length 194;
Best Local Similarity 28.1%; Pred. No. 0.092;
Matches 36; Conservative 17; Mismatches 56; Indels 19; Gaps 4;
QY 118 KQARLFEDLEDDAEVIFPSSVYPGLPESTVPANKGVASSAIQTLFTGPKVKAER---PT 174
DB 81 KKALEQTEADLKAVNEPGKAPAPAPETPAP-----EAPAEQ-----PKPAPETPAPA 130
QY 175 PTPKPKRGRKKKQADPEINSSAPRDDPENRSKFYESASARKRTVTAERERAINAA 234
DB 131 PKPEKPAEQPKPEKADQQAEDYARRSEBYNRLTQQQAPAPAKPEQPAKPEKPA---- 186
QY 235 KTFEPTNP 242
DB 187 --EETQP 192

RESULT 10
US-08-740-223A-16
Sequence 16, Application US/08740223A
Patent No. 6265564
GENERAL INFORMATION:
APPLICANT: Davis, et al.
TITLE OF INVENTION: Expressed Ligand - Vascular
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESSES: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill Road
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/740,223A
FILING DATE: 25-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/022/999
FILING DATE: 02-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Cobert, Robert J
REGISTRATION NUMBER: 36,108
REFERENCES/DOCKET NUMBER: REG 333
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

FEATURE:
NAME/KEY: htl2
LOCATION: 1...496
OTHER INFORMATION: human TIE-2 ligand 2
US-08-740-223A-16
Query Match 5.2%; Score 92.5; DB 3; Length 496;
Best Local Similarity 29.3%; Pred. No. 0.42;
Matches 27; Conservative 16; Mismatches 26; Indels 23; Gaps 4;
QY 45 GHVWRVGLRKADNKKIWFQDGHQ-----EFVDYRSIRIGYLL---IPRY 84
DB 319 GGGWTIIQRREDGSLDFQGMKVKYKFGSPSGEYWLGNFISQITNQORYVLKHLKDW 378
QY 85 EGNsafsvy-IFNLSHSEINY--HSTGLMDSA 113
DB 379 EGNEAYSlydHFYISGEELNRYIHLKGLTGTA 410

RESULT 11
US-09-709-188-16
Sequence 16, Application US/09709188
Patent No. 6441137
GENERAL INFORMATION:
APPLICANT: Davis et al.
TITLE OF INVENTION: Expressed Ligand - Vascular Intercellular Signaling Molecule
NUMBER OF SEQUENCES: 333-Z
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 08/740,223
FILING DATE: 1996-10-25
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
LENGTH: 496
TYPE: PRT
ORGANISM: Homo sapiens
US-09-709-188-16

Query Match 5.2%; Score 92.5; DB 4; Length 496;
Best Local Similarity 29.3%; Pred. No. 0.42;
Matches 27; Conservative 16; Mismatches 26; Indels 23; Gaps 4;
QY 45 GHVWRVGLRKADNKKIWFQDGHQ-----EFVDYRSIRIGYLL---IPRY 84
DB 319 GGGWTIIQRREDGSLDFQGMKVKYKFGSPSGEYWLGNFISQITNQORYVLKHLKDW 378
QY 85 EGNsafsvy-IFNLSHSEINY--HSTGLMDSA 113
DB 379 EGNEAYSlydHFYISGEELNRYIHLKGLTGTA 410

RESULT 12
US-09-270-767-44371
Sequence 44371, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 44371
LENGTH: 748
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-44371

Query Match 5.1%; Score 92; DB 4; Length 748;
Best Local Similarity 27.2%; Pred. No. 0.91;
Matches 44; Conservative 19; Mismatches 63; Indels 36; Gaps 7;

QY 98 SHSEINHYSTGLMSAHNHFKARLFDLEDEDAEVI--FPSSVYPSLPPEST-----VPA 151
Db 536 SSSILEKHLTTSRKLHHH-----SAVNDODDARVLEFANSKOPPPLAASSTTFVNA 588
QY 152 NKGVASSAIQTLFTGCPVKAEBFTPTPKIPKGRK-----KNADPEEINSSAPRDDDPEN 207
Db 589 SVGFPSA-----GSVFASSTSCSKAPKPGEEYILPDNMDVDEIVIS-----S 633
QY 208 RSKFYESASAKRTVTAE-----RERAINAKTPEPTNPPFR 245
Db 634 SSSYTSSAAQVNTLQPPDFNLFHQATTATTATYFNFGR 675

RESULT 13
US-09-919-497-98
; Sequence 98, Application US/09919497
; Patent No. 6773883
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 98
; LENGTH: 1736
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-98

Query Match 5.1%; Score 91.5; DB 4; Length 1736;
Best Local Similarity 19.4%; Pred. No. 4;
Matches 63; Conservative 41; Mismatches 126; Indels 95; Gaps 14;
QY 44 DGHVVRVGLRKA-----DNK1WFQD-----WQEFVDYRSIRIGYLLIFRYEGNSAFSVY 93
Db 768 DG--WYGAKEAVQQOQNLVWVSEKAGDAGTSDLDLHDDRLSYL-----SAPGEYSWY 821
QY 94 IFNLSHSEINHYSTGLMSAHNHFKARLFDLEDE-----DAEVIFFPS----- 138
Db 822 STDSTRHS-DYEDT---DTEGAYTDQELDELINDEVGTPPESAITRSSEFPVRESQW 877
QY 139 -----VYP-----SPLPESTVPANKGYASSAI-OTL 163
Db 878 HENOTYPPYSPQAQPIHRIDSPGFKPASQQAASSFPVYLSPETNPASTSAVNHV 937
QY 164 FTGPVKAEPTPTPKIPKGRKKKNADPEEINSSAPRDDDPENRSKFYESASAKRTVT 223
Db 938 NUTNVRLEPTPAPSTYSYQADSURTSTEAHHMLRDQEP-SLSSHVDPTKYRKDPY 996
QY 224 AEERERAINAAK-----TPEPTNPPF-----RVVLRPSYLYRCIMYL 261
Db 997 PEEMRQNHVLKQPAVSHGHRHPDKEPNLTYPQLPYVEKQASRDLEQPTRYES--SSY 1054
QY 262 PSGFAEKYLSGISGFIKIVOLAEKQW 286
Db 1055 TDQFSRNYEHLRYEDRVPMYEQW 1079

RESULT 14
US-09-762-481B-2
; Sequence 2, Application US/09762481B
; Patent No. 6632639
; GENERAL INFORMATION:
; APPLICANT: DREYFUS, MARC
; APPLICANT: LOPEZ, PASCAL
; TITLE OF INVENTION: MUTANT E. COLI STRAINS, AND THEIR USE FOR PRODUCING
; TITLE OF INVENTION: RECOMBINANT POLYPEPTIDES
; FILE REFERENCE: US9898APCNR

; CURRENT APPLICATION NUMBER: US/09/762,481B
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: PCT/FR99/01879
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: FR 98/10197
; PRIOR FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1061
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-762-481B-2

Query Match 5.0%; Score 90; DB 4; Length 1061;
Best Local Similarity 22.9%; Pred. No. 2.6;
Matches 50; Conservative 21; Mismatches 69; Indels 78; Gaps 9;
QY 93 YIFNLSHSEINHYST-----GLMDSAHNHFKARLFDL----- 126
Db 450 YLLNEKESAVNAIETRODGVRCVIVPNDQMETPHYHVLVRKGBETPTLSYMLPKLHEEA 509
QY 127 -----EDEDAEVIFPS-----SVYSPPLPES-----TVPANKG 154
Db 510 MALPSEEFABERKPEQPALATFAMPDVPFAPTPAEPAAPVVPAPAKAATPATPAAPQPG 569
QY 155 YAS---SAIOTLFTGVPKAEPTPT-----PKIPKGRKKKNADPEEINSSAPRDDDPEN 207
Db 570 LLSRFFGALKALFSG---GHETKFTEQAPAKAEAKPERQDRKPRQ--NNRDENERD 625
QY 208 RSKFYESASAK-----RTVTAERERAINAAKT 236
Db 626 RSERTEGSDNRENNRRNRQAQQQTAESTRSQQAQEV 663

RESULT 15
US-08-960-507-20
; Sequence 20, Application US/08960507
; Patent No. 6057435
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; TITLE OF INVENTION: Tie Ligands
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,507
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1130p1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-960-507-20

Query Match 4.9%; Score 88.5; DB 3; Length 286;
 Best Local Similarity 30.4%; Pred. No. 0.47;
 Matches 28; Conservative 14; Mismatches 27; Indels 23; Gaps 4;

QY 45 GHVRYGLRKADNKIWFDGQWQ-----EFVDRYSIRIGYLL---IFRY 84
 Db 109 GGGWTIIQRREDGSDVDFQRTWKYKVGFGNPGSEYWLGNFVSQLTNQORYVLKHLKDW 168

QY 85 EGNsafsvY-IFNLshSEINy--HSTGLMDSA 113
 Db 169 EGNsYSLYEHFVLSSEELNYRIHLKGLTCTA 200

Search completed: December 30, 2004, 06:15:51
 Job time : 27 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 06:10:48 ; Search time 367 Seconds
(without alignments)
334.243 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796

Sequence: 1. MPRPFHKLIFSSTIQEKRL.....LLRTRDFVLKVTAFRNEYV 341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	908.5	50.6	431	15	US-10-424-599-260947
2	554.5	30.9	188	15	US-10-424-599-240434
3	375.5	20.9	203	15	US-10-424-599-240433
4	329	18.3	344	17	US-10-739-930-9287
5	289.5	16.1	502	17	US-10-425-115-274348
6	280	15.6	375	15	US-10-425-114-63507
7	278	15.5	402	16	US-10-437-963-157411
8	266.5	14.8	181	15	US-10-424-599-166413
9	265	14.8	362	16	US-10-437-963-192114
10	251.5	14.0	750	16	US-10-437-963-178183
11	246	13.7	737	16	US-10-437-963-203869
12	241.5	13.4	322	16	US-10-437-963-152254
13	214	11.9	306	16	US-10-437-963-129669

14	210.5	11.7	164	15	US-10-424-599-206969	Sequence 206969,
15	204.5	11.4	396	17	US-10-425-115-314879	Sequence 314879,
16	197	11.0	519	16	US-10-437-963-187150	Sequence 187150,
17	195.5	10.9	1355	16	US-10-437-963-137649	Sequence 137649,
18	193.5	10.8	118	15	US-10-424-599-179398	Sequence 179398,
19	193	10.7	536	16	US-10-437-963-162720	Sequence 162720,
20	186	10.4	462	16	US-10-437-963-195577	Sequence 195577,
21	183	10.2	100	15	US-10-424-599-232448	Sequence 232448,
22	175.5	9.8	538	16	US-10-437-963-120554	Sequence 120554,
23	174	9.7	422	16	US-10-437-963-177245	Sequence 177245,
24	171.5	9.5	167	17	US-10-739-930-9695	Sequence 9695, Ap
25	170.5	9.5	391	16	US-10-437-963-138423	Sequence 138423,
26	166	9.2	270	16	US-10-437-963-156915	Sequence 156915,
27	165.5	9.2	181	15	US-10-424-599-241998	Sequence 241998,
28	162	9.0	259	15	US-10-424-599-216590	Sequence 216590,
29	161	9.0	352	15	US-10-425-114-42561	Sequence 42561, A
30	147	8.2	305	16	US-10-437-963-105662	Sequence 105662,
31	145.5	8.1	230	16	US-10-437-963-188750	Sequence 188750,
32	141	7.9	143	17	US-10-425-115-274347	Sequence 274347,
33	141	7.9	417	16	US-10-437-963-170458	Sequence 170458,
34	140.5	7.8	141	16	US-10-437-963-195574	Sequence 195574,
35	140.5	7.8	156	15	US-10-424-599-207209	Sequence 207209,
36	140	7.8	255	16	US-10-437-963-188901	Sequence 188901,
37	135.5	7.5	105	17	US-10-425-115-292035	Sequence 292035,
38	135.5	7.5	212	17	US-10-425-115-253856	Sequence 253856,
39	134	7.5	237	16	US-10-437-963-122652	Sequence 122652,
40	130	7.2	239	17	US-10-425-115-341488	Sequence 341488,
41	128.5	7.2	1070	16	US-10-437-963-167140	Sequence 167140,
42	127.5	7.1	226	17	US-10-425-115-358357	Sequence 358357,
43	126.5	7.0	156	17	US-10-425-115-221099	Sequence 221099,
44	123.5	6.9	148	15	US-10-424-599-173953	Sequence 173953,
45	122.5	6.8	87	16	US-10-767-701-54013	Sequence 54013, A

ALIGNMENTS

RESULT 1
US-10-424-599-260947
; Sequence 260947, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 260947
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_77658C.1.pap
US-10-424-599-260947

Query Match	50.6%	Score	908.5	DB	15	Length	431
Best Local Similarity	49.7%	Pred. No.	7.1e-76				
Matches	198	Conservative	46	Mismatches	95	Indels	59
Gaps	8						
QY	1	MPRPFHKLIFSSTIQ-EKRLRVPDKFVSKFKDELSSVALTVPDGHVWRVGLRKADNKI	59				
Db	1	MPHPSFKLLLPSTVQPNQQLRPLDNFMRKYGGLSPITVLSVPDGSVWVHVLKKAADNKY	60				
QY	60	WFQDCWQFVDYRYSIRIGYLLIFRYEGNSASVYIFNLSHSEINTHSTGLMDSANHFPCR	119				
Db	61	CFLDGMKEFQRYISGVGSLCVTRRVGKSVRTWIFNLATSEINYSQSVTRSSNEGLHFTN	120				
QY	120	A-RLPDEDEDAEVI---FPSSVPSPLP-----ESTVPANKGVASSAIQTLFTG-	166				

Db 121 CLKFEEMEGEDSIEISDSSPHLSPIQNALAGSVDKWMPGKSYNTPPALQNLFGS 180
QY 167 -----PVAEBEPTTPKI-----PKRG 184
Db 181 KLSINWEGGNAHSRSANSIDNRLTRDIGQFNAVEFKRSTBELKLRASIEERMKKT 240
QY 185 RKKKNADPEEINSSAPRDDPENRSKFYESASARKRTVTAERERAINAAKTFFPTNPF 244
Db 241 RKRKSDQGE--PSAGHEEVEEMRFRFYESASARKRTVTAERERAINAAKTFFPTNPF 298
QY 245 RYVLRPSYLYRGCMYLPSPGFAEKYLSGISGFIKVQLAE-KQMPVRCLYKAGRAKFSQGW 303
Db 299 RYVLRPSYLYRGCMYLPSPGFAEKHLNGVSGFIKQLISNGRQWVRCYKGRAKLSQGW 358
QY 304 YFETLENNLGEQDVCVFELLRTDRFVLKVTAFRNEVY 341
Db 359 FEFSLLENNLGEQDVCVFELLRMKEVVLQVTFPHVTEV 396

RESULT 2

US-10-424-599-240434

; Sequence 240434, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 240434

; LENGTH: 188

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_59139C.1.p

US-10-424-599-240434

Query Match 30.9%; Score 554.5; DB 15; Length 188;

Best Local Similarity 76.6%; Pred. No. 2.5e-43;

Matches 108; Conservative 14; Mismatches 14; Indels 5; Gaps 2;

QY 204 DPENRS----KFYESASARKRTVTAERERAINAAKTFFPTNPFRRVLRPSYLYRGCM 259

Db 7 DSENVADKPRWFYELASARKSTVTAERERAINAAKTFFPTNPFRCVLRPSYLYRGCM 66

QY 260 YLPSPGFAEKYLSGISGFIKVQLAE-KQMPVRCLYKAGRAKFSQGWYFETLENNLGEQDVC 318

Db 67 YLPSPGFAEKHLNGVSGFIKQLISNGRQWVRCYKGRAKLSQGWYFETLENNLGEQDVC 126

QY 319 VFELLRTDRFVLKVTAFRNE 339

Db 127 VFELLRTREVLQVTFRTE 147

RESULT 3

US-10-424-599-240433

; Sequence 240433, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 240433
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_59138C.1.p
US-10-424-599-240433

Query Match 20.9%; Score 375.5; DB 15; Length 203;

Best Local Similarity 68.5%; Pred. No. 1.4e-26;

Matches 74; Conservative 14; Mismatches 19; Indels 1; Gaps 1;

QY 182 KRGRKKKNADPEEINSSAPRDDPENRSKFYESASARKRTVTAERERAINAAKTFFPTN 241

Db 96 KTAKKRKRSEPYGEEPSGENEEAEEMRYRFYESASARKRTVTAERERAINAAKTFFPTN 155

QY 242 PFRVLRPSYLYRGCMYLPSPGFAEKYLSGISGFIKVQLAE-KQMPV 288

Db 156 PFCVLRPSYLYRGCMYLPSPGFAEKHLNGVSGFIKQLISNGRQWSV 203

RESULT 4

US-10-739-930-9287

; Sequence 9287, Application US/10739930

; Publication No. US20040216190A1

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT

; FILE REFERENCE: 38-21(53377)B

; CURRENT APPLICATION NUMBER: US/10/739,930

; CURRENT FILING DATE: 2003-12-18

; NUMBER OF SEQ ID NOS: 11088

; SEQ ID NO 9287

; LENGTH: 344

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: GLYMA-23APR03-C63260_1.p

US-10-739-930-9287

Query Match 18.3%; Score 329; DB 17; Length 344;

Best Local Similarity 27.2%; Pred. No. 6.4e-22;

Matches 94; Conservative 57; Mismatches 130; Indels 64; Gaps 10;

QY 6 FHKLIFSTIOEKRLRVPDKFVSKFDLSVAVALTPDGHVWRVGLKADNKIWFQDQW 65

Db 16 FFKIITAHNVHEGKLMIPNKVKYKRLQNTLFLKTPNGAEWKMLKKRDKIWFQKQW 75

QY 66 QEFVDRYSIRGYLLIFRYEGNSAFSVIFNLHSEINHYSTGLMDSAHNHFKAARLPED 125

Db 76 KEFAEYHSLAHGHLILFRWDVTSFQVHIFDLSALEIEY-----114

QY 126 LEDEDAEVIPPSSVVPSPLESTVPANKGVASSA1QTLFTGPVKAEEPTTPKPKRGR 185

Db 115 -----PTEI-----IKGTASNKGNESPGDEHLEC-----HRSGQ 145

QY 186 KKKKNADPEEINSSAPRDD---DPEN-----RSKFYESASARK-RTVTAERERAINAAKT 236

Db 146 KRKNSVEFLQCCQMRSKCKVKVENTMILPQALHHTATKCKGSKAMDQVLTALDRASS 205

QY 237 FEPTNPFRRVLRPSYLYRGCMYLPSPGFAEKYLS--GISGFIKQV-LAEKQMPVRC-LY 292

Db 206 FKSCNPFPLTYMHRTHISSHGSNLPMKFCRSHLDLHKRRLISLQVLSGRINWPAKYQIH 265

QY 293 KAGRA-KFSQGWYFETLENNLGEQDVCVFELLRTDRFVLKVTAFR 336

Db 266 KQKTAIRFKLSWNAFVKDNNLKVGDVCLFELVHGTKLTFVLVHIFR 310

RESULT 5

```

; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63507
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73255A02_FLI.pep
US-10-425-114-63507

Query Match 15.6%; Score 280; DB 15; Length 375;
Best Local Similarity 27.7%; Pred. No. 2.7e-17;
Matches 89; Conservative 54; Mismatches 136; Indels 42; Gaps 13;

QY 19 RLVRPDKFVSKFDKDELSSVAVALTYP-DGHVVRVGLRKADNKIWFODGWOEFVDYRISIRIG 77
DB 59 RISVPEKAFKNKQIITEEFILKSPSSAETHVGVGEKDKLFLVSGWENFAKAHLEEN 118
QY 78 YLLIFRYEGNSAFSFIYFNLSHSE--INYHSTGLMDSAHNHF-----KRARLFEDLEDEDA 131
DB 119 DLLLFAFGSGNSSFEVLVFGASGCEKVSLSFGSLGDPDMGQKQFNDVRRHGVHHSVTVS 178
QY 132 EVLFPSPVYSPL---PESTVPANKGYASSAIQTLFTGPVKABEPTPTPKIPKGRGKKX 188
DB 179 E-----DTVAPSQLVRSPNALPLKEPSGAR-----PSKYESPSSNFIVRHVATGKE 227
QY 189 NADPEEINSSAPRDDDDPENRSKFYESASARKTVTAEERERAINAAKTPEPTNPFPRVVL 248
DB 228 GTDDEVANSN-----YYISUSANKR--IGDEBKEEIIIGLA-PTRPNPVPFTLL 272
QY 249 RPSVLYR--GCIMYLPSCGFAEKYLSGISGFIKVQL--AEKQWPVRCLYKAGRAKFSQ-GW 303
DB 273 RKNHVQRNNCLI-IPSKFAADHLGERAHNIIILRPNKEKWHVSYTHSRHTCFQNLAL 331
QY 304 YEFTLENNLGEQDVCVFELLR 324
DB 332 FKPTRENKLHEGDICVPFLMK 352

RESULT 7
US-10-437-963-157411
; Sequence 157411, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 157411
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56986C.1.pep
US-10-437-963-157411

Query Match 15.5%; Score 278; DB 16; Length 402;
Best Local Similarity 26.8%; Pred. No. 4.6e-17;
Matches 103; Conservative 57; Mismatches 157; Indels 68; Gaps 16;

QY 3 RPFPHKLIFSSTIQEKRLVRPDKFV-----SKFKDELSSVAVALTVPDGHVVRVGL 52
DB 119

```

Db 18 RPHFFKVLVGDFF--KQRLKIPNFCFKHIPWESRSKAKGLKEASMAATLEGPSGRTWLVVI 75
QY 53 KADNKKIWFQDCWQBFVDRYSIRIGVLLIFRYEGNSAFSVYIFNLHSEINVHSTGLMDS 112
Db 76 RKTAGTFTTSWPKFVQDQAKLELEFVVFRVDRGNTRFTAMVFDRTACE---REDLMGG 131
QY 113 A---HNHFKRARLFDEDEDA-----EVI-----FSSVYP-----SPLPES 147
Db 132 GGGDRPKRGRPRTAASRDAARPKKDSVGKEMTYRASPGGQPLQIVDSWTFPEGS 191
QY 148 TVPANGYASSAIQILFTGPVKAEBPTTPKIPK-----KGRKKKNADPEEN- 196
Db 192 TAVKNEEADAE---LPVCELPASSAPRPHVPEGALDADGGAARGAATRSLODDLAL 247
QY 197 SSAPRDDDDPENSKVESASAKRTVTAEERERAINAAKTPTPTPFRVLRPSLYRG 256
Db 248 ASIP-----PSIRR--YKGVSRRAVATAERORATEIAHAFRSPDPCVIRMTWHVYS 301
QY 257 CIMYLPSPGFAEYL-SGISGFIKVLQAKQWVRCLYKAG-RAKFSQGWYEFTELENNLGE 314
Db 302 FMRPTGTSRQHLPRRTDVLVDPGGKVMV--LYIPNTEDRLSRGWCAPARGNCLEE 359
QY 315 GDCVCFELLRTDFVLKVTAFRNE 339
Db 360 GDCVCFELVAAAEF--RVHIFRWE 382

RESULT 8

US-10-424-599-166413
; Sequence 166413, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166413
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_121286C.1.pep
US-10-424-599-166413

Query Match 14.8%; Score 266.5; DB 15; Length 181;
Best Local Similarity 34.4%; Pred. No. 1.8e-16;
Matches 64; Conservative 33; Mismatches 66; Indels 23; Gaps 5;
QY 3 RPF-FHKLIFSSTIQEKRLRPDPKFKVSKDBLSVAVALTPDGHVWRVGLRKADNKIWF 61
Db 13 KPFHFLIITAQLQDGLKMINKFKVEKYGEGLPNALFKTPNGTWNFNLEKHDKIWF 72
QY 62 QGQWQFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSEINVHSTGLMDSA---HNHF 118
Db 73 QKGWKEFAEYHSLAHGLLVFRHGTSHQVHIFDLSSLEIDYPSKGTGKTSNHEGNK 132
QY 119 RARLFEDLEDAEVIFPSSVYPSLPSTVTPANGYASSAIQTLFTGPVKAEBPTTPK 178
Db 133 QPR-----NEENLEYLPQYQV-----RSHSKVRVENMMTLPEEAQPTDTFK 174
QY 179 IPKRG 184
Db 175 -EKSG 179

RESULT 9

US-10-437-963-192114
; Sequence 192114, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 192114
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_8836C.1.pep
US-10-437-963-192114

Query Match 14.8%; Score 265; DB 16; Length 362;
Best Local Similarity 26.6%; Pred. No. 6.4e-16;
Matches 97; Conservative 57; Mismatches 148; Indels 62; Gaps 13;
QY 20 LRPDPKFKVSKDBLSVAVALTPDGHVWRVGLRKADNKIWFQDQWQFVDRYSIRIGYL 79
Db 1 MKIPSSFNQCLQNPQTMVSLVDRSGNKSABELTSDBEGFFVHGKWEFVRDNTSQCGF 60
QY 80 LIFRYEGNSAFSVYIFNLHSEINVHSTGLMDSAHNHFKRARLFEDLEDAEVIFPSSV 139
Db 61 LVFTYDKRSQFSVTF--EPSGIDKIST---FSAHP--SKVVIKTESDEGGMVTAATT 113
QY 140 YPSLPPESTVPANK-----GYASSAIQTLFTGPV---KAEPTPTPKIPKR-----G 184
Db 114 -----EKMAPALKENNGITGKRTDVLDMEDRVVFKSSSEANVCSSRRKAGASAG 167
QY 185 RKK-----KNADPEEINSAPRDDDPENRSKPYESASAR----- 218
Db 168 KSKVTSTSHNSTRGSCSDSDNSSL--KSPNPPFLMRFLSGEVSRGRGRCVSKGQRLTVI 226
QY 219 --KRTVTAERERAINAAKTPTPTPFRVLRPSLYRGCMYLPSPGFAEYLSGISGF 276
Db 227 SQRFPVTEAEKDHQALQAREFKSNPFPAVQIMMESYVYVGFMMNIPCFVRECLPHTNKR 286
QY 277 IKVOLAE-KOMPVRCLYKAGRA--KFSQGWYEFTELENNLGEQDVCVFELLRTDFVLKVT 333
Db 287 ITLWDPOKAEVNVYVYSDRSVSGSGGKGFAGNNLEKFDVCVFELVQKDN--IKVH 344
QY 334 AFRV 337
Db 345 IYRV 348
RESULT 10
US-10-437-963-178183
; Sequence 178183, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 192114
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_8836C.1.pep
US-10-437-963-178183


```
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 178183
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_75764C.1.pap
US-10-437-963-178183

Query Match          14.0%; Score 251.5; DB 16; Length 750;
Best Local Similarity 26.9%; Pred. No. 3.2e-14;
Matches 86; Conservative 55; Mismatches 134; Indels 45; Gaps 13;

QY 22 VPDKFSVFKDELVAVALTPDGHVVRVGLAKAD-NKTIWFOGQWQEFVDRYSIRIGYLL 80
DB 424 VFARFANNFNCHISEVNLRSGETWSIGVANSOAGELVLPQWKEFVDGNGIEGDCL 483
QY 81 IFRYEG-NSAFSVYIFNLSHSEINYNHSTGLMDSAHHNFKARLFEDLEDE-DAEVIPTSS 138
DB 484 LFRYSGVSSSPDLIFDPGCE-----KASHFVSGHGFGRAGENSAGAEQGGNG 533
QY 139 VYPSLPPESTVPANKGYASSAIQTLFTGPVKAEBEPTTPKIPK--KRG-----RKKN 189
DB 534 RRTPTI-----VGDNGHRRHLEMTLHRNSCRS-----IPRACKRSLSFDETEAKEND 581
QY 190 ADPEELNSAPRDDPENRSKFVESASAR--KRTVTAERERAINAAKTPEPTNPFRV 247
DB 582 GDEDDVAAA-----EGRGVGEYFNRKHGRVAEYNLEEDREISRPVVPQGNPVQV 637
QY 248 LRPSYL--YRGIMVLPSPFABKYLGSIGSFYKVLAEK--QWPVRCLYKAG-RAKFSQG 302
DB 638 IHSSHVRSKCYCIVGSPFAGKYLGAVEREVLERASRGGEHVPFVHRQNRGFGYAG 697
QY 303 WYEFLENNLGGDVCVFL 322
DB 698 WRQFAGDNRNLVAHDVCLPEL 717

RESULT 11
US-10-437-963-203869
; Sequence 203869, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203869
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99011C.1.pap
US-10-437-963-203869

Query Match          13.7%; Score 246; DB 16; Length 737;
Best Local Similarity 23.8%; Pred. No. 1e-13;
Matches 74; Conservative 60; Mismatches 123; Indels 54; Gaps 8;

QY 5 FFHKLIFSSTIQEKLRLVPDKFVSKFDELVAVALTPDGHVVRVGLRKADNKIWFQDG 64
```

```
DB 149 YFFKVMIGGF--RRQMTIPYFAENFRDQIQGTIKLKARNGNTCSVLVDKCNKLVLTGK 206
QY 65 WQEFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINYNHSTGLMDSA-HNHFKRARLF 123
DB 207 WAEFANSHDIKMGDFLVFRYTGNSQFEVKIFDPS-----GCYKAASHNAVIGQHA 257
QY 124 EDLEDEDAEVIPTSSVYPSLPPESTVPANKGYASSAIQTLFTGPVKAEBEPTTPKIPK 183
DB 258 QNMQDPTIELSCSDHLRAQSLTTERQNO-----PEKD 291
QY 184 GRKKQVADPEEINSSAPRDDPENRSKFVESASARKTVTAEERERAINAAKTPEPTNP 243
DB 292 VIDNCNCKKXKTEHASSSEDDQ-----ETPTAEVHRMKVEMVRAIHS-----NHFP 337
QY 244 FRVLRLPSVLYR-GCIMVLPSPFABKYLGSIGSFYKVLAEKQWPVR-CLYKAGRAKFSQ 301
DB 338 FVAVWKKSNVTRQPCYVAISRYANEYFPGDQMLTLQRHGRWQVKFCISRKRLMLSK 397
QY 302 GWYEFLENNL 312
DB 398 GWRKPTRDNL 408

RESULT 12
US-10-437-963-152254
; Sequence 152254, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 152254
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_52323C.1.pap
US-10-437-963-152254

Query Match          13.4%; Score 241.5; DB 16; Length 322;
Best Local Similarity 25.2%; Pred. No. 8.5e-14;
Matches 80; Conservative 57; Mismatches 140; Indels 41; Gaps 8;

QY 18 KRLRPDPFVSKFDELVAVALTPDGHVVRVGLRKADNKIWFQDGQWQEFVDRYSIRIG 77
DB 37 ERMIIPNEFLQYFRGKIPRTIKQLRDGCTVDVQVTKLGLKISLQSGWKAFVTAHDLQNG 96
QY 78 YLLIFRYEGNSAFSVYIFNLSHSEINYNHSTGLMDSAHHNFKARLFEDLEDEDAEVIPT 137
DB 97 DFLVPSYDGI SKLVLI FPGSGCEKVSRSRTLKNATHCGEK----- 137
QY 138 SVYPSLPPESTVPANKGYASSAIQTLFTGPVKAEBEPTTPKIPKRGKKKNADPEEINS 197
DB 138 --WEEPL-----HISNSHDL-----PVKS---PQVSKSEKQWDSSEQNDTANIEE 180
QY 198 SAPRDDPENRSKFVESASARKT-VTAERERERAINAAKTPEPTNPFRVLRPSYLY-R 255
DB 181 VALQGDLLQGHVPL--NCILPKHTLTDQKQLESKVGAHSEIPIYGCILRKSRVHGK 238
QY 256 GCIMVLPSPFABKYLGSIGSFYKVLAEKQWPVRCLYKAGRAK-FSQGWYEFLENNLGE 314
```

```
Db 239 SQTVDICREYADVLPFKELNMTLQRHKGWVLCRTKDTKRLSTGWSRPAQENNLQV 298
QY 315 GDVCFVPELLRTRDFVLKV 332
Db 299 GDICLFPKELLKKEYSMMV 316

RESULT 13
US-10-437-963-129669
; Sequence 129669, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 129669
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_31904C.1.pap
US-10-437-963-129669

Query Match 11.9%; Score 214; DB 16; Length 306;
Best Local Similarity 19.7%; Pred. No. 2.9e-11;
Matches 63; Conservative 55; Mismatches 137; Indels 64; Gaps 5;

QY 22 VPDKFKSKFDELSVAVALTPDGHVWRVGLKADNKIWFQDQWQEFVDRYSIRIGYLLI 81
Db 46 IPNEFLHNFGGKIPKSIKLETRSGLTDFQVTKNSGRVVLQSGWASYSAHDLKIGDFLV 105
QY 82 FRYEGNSAFSVIFNLSHSEINHYSTGLMSAHNFKRARLFEDLEDEDAEVIFFSSVYP 141
Db 106 FKSGDSQLKTLFD----- 120
QY 142 SPLPESTVPANKGYASAIQTFTGPVKAEETPTPKPKKGRKKQGNADPEINSSAPR 201
Db 121 -----SSGCEKVCKEVDMSGRSYDIAMENSQDEKKKKRQDRISRQGTVK 165
QY 202 DDDPENRSKFEYASARKET-VTAERERAINAAKTFEPTNPFVRVLRPSLYLR-GCIM 259
Db 166 PSEEGUKAELVPGCILPSRDTLRLQKNILIEKVAINSETPIYGVVMNNSIHGIPCTV 225
QY 260 YLPSGPAEKYLSGISGFIKVLAEKQWVR-CLYKAGRAKFSQGWYFTLENNLGEQDVC 318
Db 226 EISKYADVLPEDGTVLVQHHGKSMNVRCCLTKQNSKRFKLGWRQFAGDNKHLGDI 285
QY 319 VFELLR-TRDFVLKVTAFR 336
Db 286 LFDLLKDKKKYVMDVHIIR 304

RESULT 14
US-10-424-599-206969
; Sequence 206969, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```

```
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206969
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_2891C.1.pap
US-10-424-599-206969

Query Match 11.7%; Score 210.5; DB 15; Length 164;
Best Local Similarity 42.1%; Pred. No. 2.6e-11;
Matches 45; Conservative 18; Mismatches 43; Indels 1; Gaps 1;

QY 2 PRPFHKLIFSSTIQEKRLVRPDKFVSKFDELSVAVALTPDGHVWRVGLKADNKIWF 61
Db 57 PAVHFVKIILTSLADGIL-LPKKFKTKYGDGMSNPVFLKPADGTGTEWKHIYTKHGSIWF 115
QY 62 QDGHQEFVDRYSIRIGYLLIFRYEGNSAFSVIFNLSHSEINHYSTG 108
Db 116 QKGWKEFATYYSLDHGHLLFFEYEGTSHDFDVLHFDSSALEIDIPSHG 162

RESULT 15
US-10-425-115-314879
; Sequence 314879, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 314879
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(396)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_50230C.1.pap
US-10-425-115-314879

Query Match 11.4%; Score 204.5; DB 17; Length 396;
Best Local Similarity 22.2%; Pred. No. 3.2e-10;
Matches 76; Conservative 52; Mismatches 104; Indels 11; Gaps 13;

QY 3 RPFFHKLIFSSTIQEKRLVRPDKFVSKFDE--LSVAVALTV-PDGHVWRVGLKADNKI 59
Db 136 KQOFIR-IFHNLSEK-WMLPARVKOVSEARLNSRIALIFSPIAKFWRVELKNDQSGI 193
QY 60 WFDQGWQEFVDRYSIRIGYLLIFRYEGNSAFSVIFNLSHSEINHYSTGLMSAHNFKR 119
Db 194 PFTGWSQFLDPHGKNGEVLLKVEGNWVFKFAFLSGCQ----- 235
QY 120 ARLFEDLEDEDAEVIFFPSVVPSPLESTVPANKGYASAIQTFTGPVKAEETPTPKI 179
Db 236 ----KDFKNQAGI-----QINTKKQOETPTPI 258
QY 180 PKKGRKKKNADPEE-----INSSAPRDDDDPENRSKFEYASARKETVT-AEERAIN 232
Db 259 RRRKSNDEKSSSEENKRLKTPVTSPPSPDPSLEKPYQIGTSSWIKKRIINTYALEQFLALS 318
```

[illegible]

Search completed: December 30, 2004, 06:22:52
Job time : 368 secs

THIS PAGE BLANK (USPTO)